CR 86029

HSER 4994

APOLLO APPLICATIONS PROGRAM DATA ARCHIVES

CONTRACT NAS 12-556

by

Clifford W. Williams,
George E. Schmidt, Jr., and
David L. Sharp

Submitted to

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
ELECTRONICS RESEARCH CENTER
CAMBRIDGE, MASSACHUSETTS

HAMILTON STANDARD SYSTEM CENTER
FARMINGTON, CONNECTICUT

Hamilton Standard DIVISION OF UNITED AIRCRAFT CORPORATION UA® HSER 4994

APOLLO APPLICATIONS PROGRAM DATA ARCHIVES

CONTRACT NAS 12-556

by

Clifford W. Williams,
George E. Schmidt, Jr., and
David L. Sharp

Submitted to

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
ELECTRONICS RESEARCH CENTER
CAMBRIDGE, MASSACHUSETTS

February 15, 1968

HAMILTON STANDARD SYSTEM CENTER
FARMINGTON, CONNECTICUT

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF TABLES

CONCLUSIONS AND RECOMMENDATIONS

The Apollo Applications Program (AAP) - Data Archives study has provided the opportunity to relate the work accomplished in the Mercury, Gemini and Apollo programs to the planning and development phases of the AAP. The study has reviewed and investigated those aspects of the programs which relate to collecting, storing, and distributing the experimental data and experimental support data associated with the AAP. Certain critical facts have been exposed during this study, bringing into focus the parameters to resolve the plan for establishing a repository for the experimental data. The aspect of making the AAP experimental data available to industry and other users as a public service has been addressed.

As a result, specific conclusions and/or recommendations are highlighted in this section, to direct attention to those items which are considered most significant. These items are hereby presented in an orderly, condensed format. Additional facts and data may be found by a more detailed perusal of the complete report.

The task emphasized during this period was the survey of potential users. This emphasis was necessary because the need for an archives as well as the ability to successfully interface with users had to be established as soon as possible. Only 17 percent of those contacted were aware of the AAP experiments relevant to their own scope of research, only 22 percent were aware of the results obtainable by combining data from AAP sensors and none were aware that data from the AAP would be made publicly available. Most of the potential users contacted are eager for data and already have equipment capable of handling it. The eagerness has demonstrated itself in an expressed desire by some companies to design and submit their own experiments and to participate in a pilot Apollo Applications Program Archives System. This study was hampered by a lack of open literature from NASA concerning the approved AAP experiments and by a general lack of publicity concerning the AAP itself. Thus, it became necessary to devote a great deal of time to educating potential users. The study has shown that if appropriate information is made available, the archives will be replenished by new data from experiments generated by industrial and academic communities. The new experiments will be a direct result of "in depth" investigations into experimental data made available from the archives.

A continuation of the survey is recommended if NASA expects to work with industry toward uses of experimental data derived from AAP. The highly organized method of operation derived during this contract has provided the key elements (i.e., initial contact point, prebriefing letter presentation, continued liaison) for an operating nucleus which can be expanded. In the event that it is impossible to continue the liaison established during this study, action should be taken to dissolve this relationship to prevent loss of confidence in the ability of NASA to deliver required information to the potential user.

As it has been a key element in this study, the interface with users will also be an essential feature in the development of the Apollo Application Program Archives (AAPA). For an undertaking such as the AAPA, it is extremely important to have an active body of users who are both aware of the system and the data stored therein, as well as the mechanics for extracting or retrieving the data. Thus, it is recommended that the user visitations and presentations be continued on an even larger scale to a broader base of industrial, academic, and governmental representatives. Experience gained from this study further indicates that in order to accommodate the small business activities with the AAPA, it will be necessary to provide a central repository with some peripheral equipment for analyzing data.

With so many experiments having a potentially large user interest, it would be possible to establish the pilot AAPA on data from AAP 1-A alone. The data itself is of extremely diverse content and in formats so various and numerous that the types of record media to be handled are not even known in advance. Data use is by a geographically dispersed group of users who in total cannot be identified in advance with certainty and who have no organization, subject or discipline in common. It is a messy situation and one which is potentially quite expensive, or disastrous, or both for the AAPA. Effective steps to manage such a difficult situation will require immediate action to create a broad and facile information tool to support the AAPA data control subsystem.

The first thing required in any further investigation is a rigid definition of the scope of the system; i.e., what direction will the system take, what data will it store, and what data classes (experimenter tapes, reduced data tapes, raw data tapes, etc.) will it store. With this completed, the functional details behind the first level flow diagram (Section 3.7) may be established.

It is believed that at least initially, the AAPA should be a self-contained system storing all non-space-oriented AAP data (the space-oriented data to be stored within NASA at the NSSDC). As such it will have its own distribution scheme. Since the system will service government, industry, and the scholastic community, the detailed nature of the distribution mechanism will accordingly depend both upon the users' facilities and the frequency with which each user queries the archive.

The initial system, aimed at fulfilling the system requirements through a limited time period (about two to three years), will provide a basic working system, conservatively outfitted but capable of multidirectional expansion. While the system is evolving, additional user information can be gathered, and as this data accrues, the systematic expansion can then be planned to enable a closer fit between system demands and performance capabilities. This approach is in keeping with the general economics of a system such as the AAPA. Since the utilization requirements will be small at first and grow with time, there is no justification for implementing a full-blown system at the outset when it would not be utilized.

It has been concluded that although user requirements must be given due consideration in determining which experiments data are to be stored in the AAPA, the exact requirements cannot be predicted in advance with a full degree of accuracy. Nor is across-the-board active storage of all data the solution; this operates at cross purposes with the effective and economic retrieval of data required to satisfy user needs. The argument then, is strongly in favor of a dynamic system involving the concepts of both active and inactive (or less active) storage, and depending on the retirement of low usage data for preservation of effective retrieval. This is the basis of the system concept outlined in Section 3.8.

The requirements of the users of the experimental data do not, at this time, justify a decentralized system of distribution. The types and forms of experimental data that will be available for distribution, according to the present experiment listing, do not require real time or near real time distribution. The number and location of potential users presently interested in the experimental data do require a source of information concerning AAP experiments, but during the initial phase of the program, the demand for experimental data is unlikely to be so great as to justify the increased facilities.

NSSDC is the logical repository for all space science data resulting from AAP experiments. Because of the fact that the information is applied to different disciplines and that the users will probably require different types of information and have differing needs, it is recommended that the NSSDC and the AAP Data Archives should be separate organizations. Up to thirty-five percent of the presently scheduled AAP experiments are space science oriented and should be housed at NSSDC.

It is clear that there is neither a complex of software, nor one piece of hardware, which can perform all of the major functions involved in a total information storage and retrieval system. Such a system requires pieces of hardware and software tied together by a set of procedures. If the AAPA is to function principally as a straightforward information retrieval system, the business-oriented computers seem operationally adequate. If the archive operation develops more toward retrieval and specialized output processing to suit individual users' needs, it can be generally assumed that this will constitute an additional need for scientific processing capability. In this case, the procurement of main frame equipment more oriented towards scientific processing may be justified. It was ascertained that state-of-the-art equipment and software is adequate for a pilot system.

Two outstanding problems are 1) the need to educate the broadest possible base of potential users and 2) the need to provide each user with sufficient information about the background of experiments to permit a self-analysis of data requirements. To help alleviate these problems, while at the same time providing a positive step toward the creation and organization of the AAPA, it is further recommended that NASA undertake to review and categorize (under S.I.C. sector number designations) each experiment. That is, at some time during the cycle of approving an experiment, each experiment should be reviewed to determine its applicability to an economic (as opposed to scientific) endeavor and then an effort made to pinpoint the specific economic sector. This would compliment NASA's internal bookkeeping, initialize an AAPA reference index and provide for efficient distribution of information to potential users.

SUMMARY

The Apollo Applications Program Data Archives (AAPA) study was made for the purpose of identifying the requirements for a system to handle the experiments-related data associated with this program. The work performed consisted of the five tasks listed below:

1) Survey of field of potential users to determine what their requirements will be

2) Definition of performance specifications and development of a first level flow chart for information storage and retrieval system to serve the needs of future inquiries to the archives

3) Study of alternative structures which the archives might assume and recommendations regarding the amount of decentralization required

4) Study of the feasibility of including the archives within, or joining them to, the National Space Science Data Center (NSSDC) at the Goddard Space Flight Center

5) Assessment of the state of the art of ADP hardware particularly suited to the unique information processing and storage tasks associated with a data archive

The period of performance for this project was from May 15, 1967 until February 15, 1968. The results of the performance of these five tasks are contained in this report. The task emphasized during this period was the survey of potential users. This emphasis was necessary because the need for such archives as well as their ability to successfully interface with their users had to be established as soon as possible.

Following is a list of the highlights accomplished in the performance of the five tasks:

- Research Highlight: Designed and Developed Unique User/Space Experiment Correlation Technique

- Analysis of Experiment Data to be Made Available

- Formal User Briefing Designed and Developed

- Review of NSSDC; Existing System and Charter

  User Survey Conducted. 32 Government Agencies Contacted

15 Presentations Made to Industrial Firms

3 Meetings With Academic Institutions

- Established AAPA Performance Boundaries and Assembled Criteria of Operation
- Estimate of Data Base Established
- Analysis of Distribution Methods Based Upon Preliminary Requirements
- Review of Software Consideration for Library Forms
- Established a Preliminary List of Primary and Secondary Input Sources to the Archives
- Investigation and Analysis of AAPA Input/Output Interface Requirements
- Developed Factors for Future Growth Pattern of Archives
- Preliminary Throughput Analysis of AAPA System Conducted
- Conducted State-of-the-Art Assessment
- Preparation of Bibliography for AAP Archives and Data Flow
- AAPA System Specifications Defined for "Pilot" and "Operational" Phase
- Development of AAPA System Functions and "First Level" Flow Chart
- Detailed Analysis Performed on Critical Control System for AAPA
- Generated Staffing Requirements for AAPA
- Potential User's Requirements Analyzed and Discussed
- Review of AAPA System Specifications for AAP Flight IA
- Developed and Analyzed Impact of Geophraphical Distribution of Users on AAPA Design

APOLLO APPLICATIONS PROGRAM DATA ARCHIVES

By Clifford W. Williams, George E. Schmidt, Jr., and David L. Sharp
Hamilton Standard System Center
Farmington, Connecticut

1.0 INTRODUCTION

The primary objective of the AAP Experiments Data Archives study has been to identify and investigate the requirements for a system that will collect, store, retrieve, and distribute experiments-related data. The study of the system has been oriented towards the user of the data because a long-term need must be established before implementation begins to justify the expenditures to relate the experimental data of AAP to existing scientific and engineering facts. It is understood that the archives will provide a public service function and will tend not only to pull together technical information from different agencies of the federal government but to afford industry a new frontier of endeavor so as to increase their profitability. Furthermore, the study has shown that if appropriate information is made available, the archives will be replenished by new data from experiments generated by industrial and academic communities. The new experiments will be a direct result of "in depth" investigations into experimental data made available from the archives.

This study has attempted to bring into focus the facts surrounding and related to the establishment of such an archive. Consideration has been given to the present posture of the space programs, including the acquisition, communication, and reduction capabilities of NASA and the major types of data and information (i.e., scientific, operational, intelligence, and experimental) that is presently being released. It was also necessary to examine historical precedence in data handling and factor this into a practical plan of implementation for the AAP archives. The study was divided into five separate tasks. These tasks were found to be interrelated and interdependent to some degree. Figure 1 (Study Tasks) identifies these tasks and their relationships. During the initial phase of the study it was determined that Task 4 (Assessment of Present Archival Facility), Task 5 (ADP Hardware State of Art Assessment) and Task 1 (User Requirements) had to be partially completed before Task 2 (Performance Specifications) and Task 3 (Systems Requirements) could be started. Information, made available by NASA at the outset of the study, substantiated this course of action. At the outset equal emphasis was placed upon the performance of each task.

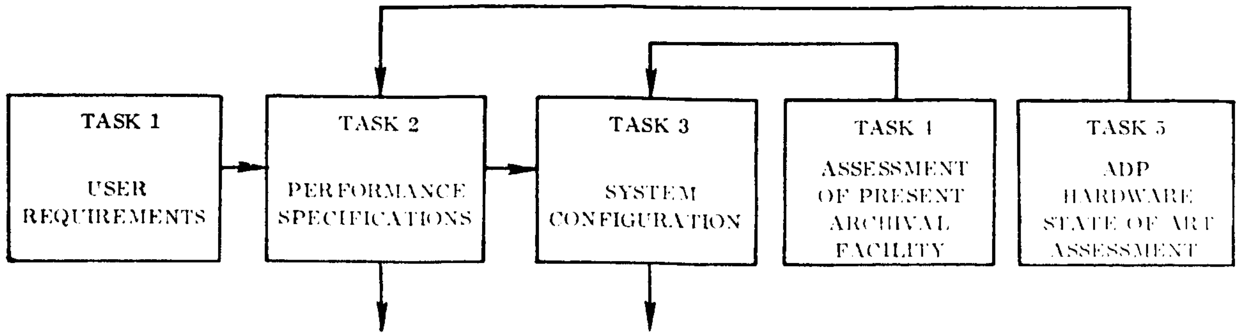


Figure 1 Study Tasks

After reviewing the descriptive literature for each experiment, initial discussions were held with potential users of the experimental data. During these discussions and following the review of the National Space Science Data Center (NSSDC) Visitors Log, additional emphasis was placed upon Task I (User Requirements) because of the need to extricate and prepare potential users before obtaining responses to a survey questionnaire. The information transferred to users was collected, sorted, selected, and formulated into prebriefing and presentation packages. This activity was completed with the aid of NASA/ERC so that data requirements of potential users could be obtained. Although results were obtained from a relatively few major organizations, it can be shown that these results represent rather large segments of the economy. Each industry briefing audience contained an average of 15 people.

The investigation of the AAP experiments-related data descriptions in documentation furnished by NASA/ERC formed a foundation upon which the state of the art assessment (Task 5) and the study of the present facility at NSSDC (Task 4) could be made. As the study progressed in these areas, it was ascertained that state of the-art equipment and software is adequate for a pilot system. Furthermore, it was found during the assessment of NSSDC that this repository would not serve the total interests of the AAP. This investigation has shown that other topical or discipline-oriented repositories, as well as NSSDC are, in fact, additional sources of experimental data and information.

As a result of these efforts, (i.e., Tasks 1, 4 and 5) a performance specification and system configuration (i.e., Tasks 2 and 3) has been established. The performance specification considers the size and type of data base, software, distribution methods, personnel, and cost considerations. A considerable amount of time was spent during this study to define specific terms since it was found that communications improved between NASA-oriented groups and other organizations when this was accomplished. A first level flow chart of the operation of the archives was developed after relating such factors as geographical acquisition and distribution, decentralization, present charter mandates, the presently defined AAP data flow system, and precedence of procedures already established. As many operational characteristics were investigated as time would allow, and functional specifications were outlined for an initial system. The functional specifications considered the necessary interfaces and centralization of operation as required.

To assure continuation of this effort, an implementation plan which includes a pilot system with the capability for expansion into the ultimate system is recommended. The further definition of users' requirements is desirable as pointed out in the user's requirements critique which was delivered during the course of this study. Further investigations should be made into data acquisition sources (i.e., ground truth data) and liaisons should be established with secondary sources as users requirements might dictate. Additional effort should be directed towards furhter defining equipment and software for record keeping. As the plan for AAP becomes more definite, projections should be made to maintain and to strengthen the archive and distribution system of the AAPA.

2.0 USER REQUIREMENTS

A key to the success of such a study (to determine the requirements for systems research of a data storage and retrieval facility) is an accurate assessment of demands of the system. The systems specifications are dependent not only upon identifying the experiments to be performed and the locations within the Federal Government by and for whom they will be performed, but also upon identifying the potential user of the experimental and engineering data. Since the experiments provide experimental data for a wide range of scientific and engineering disciplines, the potential users of this information in the academic and industrial environment appear to be abundant. Therefore it has been necessary to judiciously choose representatives of various industries in a manner to give meaning to the assessment of the specification. To a large extent, these specifications are contingent upon determining what the needs and desires of the users will be. Thus considerable time and effort has been expended to determine these user needs. The avenue of approach which has been followed is to contact potential users directly. In addition to this assessment, it has been necessary to investigate existing data storage centers such as the National Space Science Data Center (NSSDC) and the regional dissemination centers sponsored by NASA's Technology Utilization. While both of these mechanisms have proven informative, the information which comes directly from the potential users will determine feasibility, size, interface requirements usage and deployment. Presented in this section is a review of the techniques employed for estimating user requirements and of a scheme for correlating potential users with planned experiments, and a discussion of the results achieved.

2.1 Techniques for Estimating Requirements

An archive or distribution center must have a source or input of material to store or distribute to a user of the facility. The demand or pressure put upon the storage and distribution function is created by the user. Therefore any technique applied to estimate the requirements of a storage and distribution facility must first determine the needs of the user. The user must know what material is available to him if he is to specify his requirements and must relate the material of the archive or distribution system to his interests. In addition to the direct determination of user interests, the source of the material coming into the archive must be determined. The techniques for estimating requirements of a data archive are therefore bounded by the constraints of the input and output parameters.

Any one of several techniques could be applied to the Apollo Applications Program (AAP) Data Archives. The fact that precedence has been established on other NASA programs, that facilities such as NSSDC are operational, and that some AAP Experiments and Integration studies have been initiated limits the choices of techniques for estimating the requirements. Furthermore the implementation of the archives will be constrained by choosing "state of the art" equipment which must be in concert with the potential users' equipment. Other data sources for the potential user such as ESSA and the Geodedic Survey and Navy Oceanography libraries must be considered either as a companion source or an input to the Data Archives.

Considering these and other aspects of the known techniques it was decided to utilize the correlation scheme described below to identify potential users and to interrogate these potential users as to their requirements. This study, however, was hampered by a lack of open literature from NASA concerning the approved AAP experiments and by a general lack of publicity concerning the AAP itself. Thus it became necessary to devote a great deal of time to educating potential users and bring them to a point where they are now in a position to make a reasonable decision about their usage of an AAP data archives.

2.2 Correlation Scheme for Potential Users and Experiments

In order to identify the potential users, a means was devised of associating experiments and potential industrial users via the Standard Industrial Classification (S.I.C.) codes. This technique, submitted to NASA/ERC as a research highlight and described in the First Quarterly Progress Report (pp. 3-16), has thus far proven very successful and has paved the way for some very fruitful discussions.

In summary, the technique separates industry into the eighty-one industrial sectors of the 1966 Scientific American economic input/output chart (q.v., "Input/Output Economics", prepared by the editors of Scientific American, 1966). For convenience these sectors can be further grouped under nine major categories: Resource Oriented Industry, Medicine and Medically Oriented, Communications, Optics and Instrumentation, Metal Product Manufacturing, Agricultural, Non-metal Manufacturing, Research and Development, and Services and Miscellaneous. These nine groupings were revised from the seven used by Scientific American to better accommodate the purposes of the AAPA and to provide greater compatibility with the probable areas of space experimentation. Utilizing S.I.C. designations, industrial firms were listed by sector numbers and then cross-correlated with the experiments. While the assignment of experiments was somewhat more arbitrary, the method was very successful. A list of the eighty-one sectors, by group, follows in Table I.

Table II which follows is a selective list of experiments that are planned for the AAP. An attempt has been made to project an obvious usage and to assign economy sector numbers to each. Table III provides a list of potential archive users which includes most of the segments of industry that can be related to experiments in the Apollo Applications Program.

TABLE I
GROUPING OF INDUSTRIAL SECTORS

A. RESOURCE ORIENTED

Sector No.	Industry Sector	S.I.C. Nos.
37	Nonferrous Metal Ores Mining	102, 103, 104 105, 108, 109
38	Primary Iron and Steel Manufacturing	331, 332, 3391, 3399
39	Iron and Ferroalloy Ores Mining	101, 106
41	Stone and Clay Mining and Manufacturing	145, 148, 149
43	Glass and Glass Products	321, 322, 323
45	Paper and Allied Products, Except Containers	261, 262, 263 264, 266
47	Lumber and Wood Products, Except Containers	241, 242, 243 249
48	Forestry and Fishery Products	074, 081, 082 084, 086, 091
50	Rubber and Miscellaneous Products	301, 302, 303 306, 307
54	Livestock and Livestock Products	013, 014, 0193 02, 0729
55	Misc. Agricultural Products	011, 012, 0192 0199, 014, 02
56	Agriculture, Forestry and Fishery Services	071, 0723, 085 098, 0729
57	Plastics and Synthetic Materials	282
58	Chemicals and Selected Chemical Products	281, 286, 287 289

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
59	Chemical and Fertilizer Mineral Mining	147
60	Petroleum Refining and Related Industries	291, 295, 299
61	Electricity, Gas and Water	49
62	Coal Mining	11, 12
63	Crude Petroleum and Natural Gass	1311, 1321
64	Federal Government Enterprises	None Applicable
66	State and Local Government Enterprises	None Applicable
73	Finance and Insurance	60, 61, 62, 63, 64, 66, 67
78	Research and Development*	None Applicable
B. MEDICAL		
7	Drugs, Cleaning and Toilet Preparations	283, 284
71	Medical and Educational Services, Nonprofit Organizations	0722, 7361, 80, 82, 84, 86, 8921
C. COMMUNICATIONS		
13	Radio, T.V. and Communication Equipment	365, 366
69	Radio and T.V. Broadcasting	483
74	Communications, Except Radio and T.V. Broadcasting	481, 482, 489

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
D. OPTICS AND INSTRUMENTATION		
16	Optical, Ophthalmic and Photographic Equipment	381, 385, 386
19	Scientific and Controlling Instruments	381, 382, 384 387
E. RESEARCH AND DEVELOPMENT		
78	Research and Development*	None Applicable
F. METAL PRODUCT MANUFACTURING		
9	Special Industry Machinery and Equipment	3551, 3552, 3553,
10	Ordnance and Accessories	191, 1925, 1929, 193, 194, 195, 196, 199
11	Aircraft and Parts	372
12	Miscellaneous Transportation Equipment	373, 374, 375, 379
14	Materials - Handling Machinery and Equipment	3534, 3535, 3536, 3537
15	Misc. Manufacturing	391, 393, 394, 395, 396, 398/9
17	Service-Industry Machines	3581, 3582, 3584, 3585, 3586, 3589
18	Household Appliances	363
20	Office, Computing and Accounting	357

* Includes colleges and universities

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
F. METAL PRODUCT MANUFACTURING (continued)		
21	Farm Machinery and Equipment	352
22	Engines and Turbines	351
23	Construction, Mining and Oil Field Machinery	3531, 3532, 3533
24	Misc. Electrical Machinery, Equipment and Supplies	369
25	Metalworking Machinery and Equipment	354
26	Motor Vehicles and Equipment	371
27	General Industrial Machinery and Equipment	364
28	Electric-Lighting and Wiring Equipment	364
29	Electrical Industrial Equipment and Apparatus	361, 362
30	Electronic Components and Accessories	367
31	Heating, Plumbing and Structural Metal Products	343, 344
32	Machine Shop Products	359
33	Metal Containers	3411, 3491
34	Stampings, Screw Machine Products and Bolts	345, 346

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
F. METAL PRODUCT MANUFACTURING (continued)		
35	Other Fabricated Metal Products	342, 347, 348, 349
36	Primary Nonferrous Metal Manufacturing	2819, 333, 334, 335, 336, 3392
G. AGRICULTURAL USERS		
4	Tobacco Manufacturers	211, 212, 213, 214
5	Apparel	225, 231, 232, 233, 234, 234, 236, 237, 238
6	Misc. Fabricated Textile Preparations	239
8	Food and Kindred Products	201, 202, 203, 204, 205, 206, 207, 208, 209
51	Borad and Narrow Fabrics, Yarn and Thread Mills	221, 222, 223, 224, 225, 226, 228
H. NONMETAL MANUFACTURING		
1	Footwear and Other Leather	313, 314, 315, 316, 317, 319
2	Misc. Furniture	252, 253, 254, 259
3	Household Furniture	251
40	Stone and Clay Products	324, 325, 326, 327, 328, 329
42	Printing and Publishing	271, 272, 273, 274, 275, 276, 277, 278, 279

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
H. NONMETAL MANUFACTURING (continued)		
44	Paperboard Containers and Boxes	265
46	Wooden Containers	244
49	Misc. Textile Goods and Floor Coverings	227, 229
52	Paints and Allied Products	285
53	Leather Tanning and Industrial Leather Products	311, 312
I. SERVICES AND MISCELLANEOUS		
64	Federal Government Enterprises	None Applicable
65	Transportation and Warehousing	40, 41, 42, 44, 45, 46, 47
66	State and Local Government Enterprises	None Applicable
67	Hotels; Personal and Repair Services, Except Automobile	70, 72, 76
68	Automobile Repair and Services	75
70	Amusements	78, 79
72	Wholesale and Retail Trade	50, 52, 53, 54, 55, 56, 57, 58, 59
75	Business Services	73, 81, 89
76	Real Estate and Rental	65
77	Misc. and Repair Construction	15, 16, 17

TABLE I (continued)

Sector No.	Industry Sector	S.I.C. Nos.
I. SERVICES AND MISCELLANEOUS (continued)		
79	Office Supplies	None Applicable
80	Business Travel, Entertainment and Gifts	None Applicable
81	Scrap; Used and Secondhand Goods	None Applicable

TABLE II
SELECTIVE EXPERIMENTS (Sheet 1 of 5)

Experiment Number	Title	Application	Related Sector Numbers
T020	Jet-Shoes	Obtain engineering and design data to determine the feasibility of the "Jet-Shoe" concept.	58, 78, 16, 19
T017	Meteoroid Impact and Erosion	Determine the effects of meteoroid erosion on glass including cratering flux, erosion rate, optical properties, etc.	16, 43, 78, 64
T004	Frog Otolith Function	Recording of the otolith responses in a weightless state to relate to sensory and motor disorders.	71, 78, 19, 30
M053	Human Vestibular Function	Study the vestibular functions in a weightless state and the effects of weightlessness themselves.	71, 78, 64
M050	Metabolic Activity	Comparison of the metabolic activity in performing similar tasks both under normal and zero g conditions.	71, 78, 19, 64
M018	Vectocardiogram	Investigate effects on the heart caused by working under a null gravity environment.	71, 78, 19, 7, 64

TABLE II
SELECTIVE EXPERIMENTS (Sheet 2 of 5)

Experiment Number	Title	Application	Related Sector Numbers
M493	Electron Beam Welding	Determine performance parameters and feasibility of electron beam welding under zero gravity and vacuum conditions.	78, 24, 29, 64
M494	High-Pressure Gas Expulsion	Investigate the effectiveness of internal fins within the bottle for increasing the amount of gas delivered, etc.	78, 11, 31, 64
M479	Zero Gravity Flammability	Investigate how a flame will propagate and means of extinguishing a flame under zero g conditions	78, 58, 64
M439	Star-Horizon Automatic Tracking	Space navigation and guidance	78, 30, 11, 16, 19, 64
M423	Hydrostatic Gas Bearing	Gyroscope development	78, 30, 11, 19, 64
M466	Space Suit Evaluation	Evaluation of all phases of space suit performance	58, 57, 50, 19, 78, 5, 64
M401	Lunar Mapping and Survey System	Potential geological evaluation of the lunar surface	37, 39, 41, 59, 64, 73, 78, 16

TABLE II
SELECTIVE EXPERIMENTS (Sheet 3 of 5)

Experiment Number	Title	Application	Related Sector Numbers
S005	Synoptic Terrain Photography	Help underdeveloped countries discover and use natural resources such as oil, coal, and ground water.	37, 39, 41, 48, 56, 59, 61, 62, 63, 64, 73, 78,
S006	Synoptic Weather Photography	Provide more detailed weather photographs than those provided by the weather satellites.	78, 64, 66, 16
S039	Day-Night Camera System	Provide cloud cover images which may be correlated with other meteorological experiments; provide data on night glow background on cloud cover imaging under no moon condition; and provide experience to assess the usefulness of the image orthrion camera as a highly sensitive imaging system.	78, 64, 66, 16, 19, 13
S040	Dielectric Tape Camera	Provide correlative weather data; support disciplines of agriculture, geology, and cartography with high resolution television; provide data to assess the usefulness of a tape camera as a general purpose high resolution exploratory device.	37, 39, 41, 56, 48, 59, 62, 63, 64, 66, 78, 13, 16, 19, 73

TABLE II

SELECTIVE EXPERIMENTS (Sheet 4 of 5)

Experiment Number	Title	Application	Related Sector Numbers
S018	Micrometeorite Collection	Collect small micrometeorites and provide flux data for larger micrometeorites; and provide biological micrometeorite exposure data.	78, 71, 7, 16, 19
S015	Zero G Single Human Cells	Results of prolonged exposure of human cells to a zero gravity environment.	78, 71, 7, 16, 19
S009	Nuclear Emulsion	Yield information on the propagation and interaction of cosmic rays in space.	78, 71, 13, 69, 74, 19
S071	Circadian Rythm-Pocket Mice	Study the effects of prolonged space flight on the circadian rythm of a mammalian system.	7, 71, 78, 19
S070	UV/X-Ray Photography	Obtain data on the solar spectrum to aid in instrument design for AOSO and provide prediction data on solar flares.	13, 69, 74, 78, 16, 19
S065	Multi-Band Terrain Photography	Obtain additional information and experience relative to the fields of optics and spectral reflectance.	16, 19, 64, 78

TABLE II

SELECTIVE EXPERIMENTS (Sheet 5 of 5)

Experiment Number	Title	Application	Related Sector Numbers
S061	Potato Respiration	Biorhythmicity	19, 71, 78, 64
S050	15 Micron Grating Spectrometer	Determine atmospheric temperature profiles for meteorological applications.	64, 19, 30, 65
S048	UHF Spherics Detection	Meteorological	64, 19, 30

TABLE III

POTENTIAL ARCHIVE USERS (Page 1 of 4)

INDUSTRIAL FIRM	SIC NUMBERS	SECTOR NUMBERS
American Telephone and Telegraph Company 195 Broadway New York, New York 10017	4811, 4812	74
American Electric Power Company, Incorporated 2 Broadway New York, New York	8931, 8911, 8111, 7392	75
American Mineral Spirits Company (Div. Union Oil Co. of Calif.) Mountain Avenue Murray Hill, New Jersey 07971	2911, 2818, 2819	58, 60
Ampex Corporation 401 Broadway Redwood, California 94063	3651, 3811	13, 19
Armour and Company 401 N. Wabash Avenue Chicago, Illinois 60611	2096, 2042, 2872, 2841, 2818, 2011, 2834, 3531, 3629	7, 8, 23, 29, 58
Central Appalachian Coal Co. (American Electric Power Co. System) Montgomery, West Virginia 25136	1211	62
Chase Manhattan Bank 1 Chase Manhattan Plaza New York, New York 10015	6025	73

TABLE III
POTENTIAL ARCHIVE USERS (Page 2 of 4)

INDUSTRIAL FIRM	SIC NUMBERS	SECTOR NUMBERS
Clevite Corporation 17000 St. and Clair Avenue Cleveland, Ohio	2714, 3069, 3399, 3611, 3497, 1929, 3566, 3679	10, 26, 29, 30, 35, 38 50
Connecticut General Life Insurance Company Hartford, Connecticut 06115	6312	73
Consolidated Natural Gas Co. 30 Rockefeller Plaza New York, New York 10020	4922	61
E.I. duPont de Nemours and Co. DuPont Building Wilmington, Delaware 19801	2892, 281, 282, 2851, 3069, 3029, 2899, 2879	50, 52, 57, 58
Eastman Kodak Company 343 State Street Rochester, New York	3861	16
W. R. Grace and Company 7 Hanover Square New York, New York 10005	281, 2879, 4511, 3079, 2071, 2042	8, 50, 58, 65
Georgia Pacific Corporation Equitable Building Portland, Oregon 98204	2432, 2641, 2621, 2421 2661, 2899, 3275	40, 45, 47, 58
International Business Machines Corporation Armonk, New York 10504	1941, 357, 2645	10, 20, 45

TABLE III
POTENTIAL ARCHIVE USERS (Page 3 of 4)

INDUSTRIAL FIRM	SIC NUMBERS	SECTOR NUMBERS
International Nickel Company, Inc. 67 Wall Street New York, New York 10005	3356	36
Kennecott Copper Corporation 161 East 42nd Street New York, New York 10017	1021, 1069, 104, 1031	37, 39
Kollsman Instrument Corporation (Subs. Standard Kollsman Industries, Inc.) 80-08 45th Avenue Elmhurst, New York 11373	3811, 3921, 3831 3611, 3673, 3662	13, 16, 19, 29, 30
Phillips Petroleum Company Phillips Building Bartlesville, Oklahoma	1311	63
Radio Corporation of America 30 Rockefeller Plaza New York, New York 10020	3679, 4833, 4899	13, 30, 69, 74
Shell Oil Company 50 W. 50th Street New York, New York 10020	281, 282, 2879, 2999, 1311	58, 60, 63
Texas Gulf Sulfur Company 200 Park Avenue New York, New York 10017	1477, 1311, 281	58, 59, 63

TABLE III

POTENTIAL ARCHIVE USERS (Page 4 of 4)

INDUSTRIAL FIRM	SIC NUMBERS	SECTOR NUMBERS
United States Steel Corporation 71 Boradway New York, New York 10006	3312	38
Weyerhaeuser Company Tacoma, Washington 98401	2631, 242, 243, 2499 2621	45, 47

2.3 Description of the Survey

Early in the program it was recognized that some type of presentation would be necessary for describing the AAP to potential users of the data storage and retrieval system. The presentation was designed not only to present examples of space acquired data but also to show some of the use to which this data can be applied. The philosophy of the program and its objectives were explained in terms of the program's goals so that a solicitation could be made for estimates of projected requirements.

As a background for the presentations, a study was made of sensors with potential application to Earth resources. While there are only a few experiments to be performed in this area, if was felt that the concept of exploitation of Earth resources from space would stimulate much interest both for the presentation and also in the actual AAPA operation. Two tables which summarize the results of the study are presented below. The first, Table IV, was composed by Peter Badgley of NASA and R.J.P. Lyon of ARL to demonstrate the characteristics of the electromagnetic spectrum which are suitable for terrestrial applications. The table includes such things as data type and storage formats of representative sensors in each frequency range of the spectrum as well as some of the possible applications. The second table, Table V, lists a group of instruments and sensors currently being tested for use in space and indicates the potential applications of each.

In order to obtain maximum communication with a selected group, the presentation was divided into two parts. The first part consisted of an introductory package containing a letter of transmittal, three colored space photographs, a description of three representative AAP experiments, a booklet describing the AAP, and a questionnaire on items pertaining to the Data Archives for discussion. The second part consisted of a visitation to the group by a NASA and United Aircraft representative for the purpose of an oral presentation and an assessment of the initial interest.

Below, a brief discussion of the survey is provided. This discussion is basically a review, as a complete report (<u>Critique on User Requirements</u>) has already been submitted to NASA.

2.3.1 <u>Users Interrogated</u>. - In the course of the survey more than twenty presentations were made to industrial concerns, scientists (mostly principle investigators of approved AAP experiments), universities, and government agencies. An attempt was made to involve only the top corporate management in industry to insure the most knowledgeable and authoritative responses and to evoke the fastest possible action. The list of presentations is plotted in Figure 2 according to the data rendered.

2.3.2 <u>Response to the Briefing</u>. - The briefings to top level corporate management proved very successful though it was found that few people in industry were aware of the potential applications of Earth-oriented space data, let alone of the AAP. Thus

TABLE IV

CHARACTERISTICS OF THE ELECTROMAGNETIC SPECTRUM WHICH APPEAR USEFUL FOR TERRESTRIAL APPLICATIONS

						Visible			Infrared				
	γ-rays	Hard X-rays	Soft X-rays	Vacuum Ultraviolet	Near UV	Photog UV	Spectral Vis	Photog IR	Near Infrared	Med. Infrared	Far Infrared	Microwave	Radar
Frequency	10^{14} mHz	10^{13} mHz	10^{12} mHz	3×10^{10} to 10^{9} mHz	8×10^{8} mHz	Around 5×10^{8} mHz			About 10^{3} mHz	About 10^{7} mHz	About 10^{6} mHz	---	20,000 to 300mHz
Wavelength	0.03 Å	0.3 Å	3 to 100 Å	100 to 3000 Å	Near 4,000 Å	4,000 to 5,000 Å	5,000 to 7,000 A	7,000 to 10,000 Å	0.7 to 4 microns	4 to 15 microns	15 to 800 microns	1 mm to 100 cm	K-Band 1.5 cm to P-Band 100 cm
Energy	1×10^{6} ev	5×10^{4} ev	1×10^{3} ev	12 ev to 4 ev	3 ev	2.7 to 1.75 ev		1.75 to 1.25 ev	1.75 to 0.3 ev	0.3 to 0.08 ev	Very low	Very low	
Operational Mode	Passive-scintillation counters			Active (2,000 to 3,900 Å)	Passive	Generally passive			Reflectance (solar-active)	Emission (passive)	Emission (passive)	Passive, operated tuned to one wavelength	Active, CW or pulsed, tuned to one wavelength
Atmosphere attenuation	Extreme attenuation			2,200 to 2,700 Å, Rayleigh scattering; 2,400 to 3,100 Å, ozone strongly absorbs	Limited attenuation	Limited attenua-tion	Good visi-bility	Better visibility	Extreme in bands at 0.9, 1.13, 1.38, 1.9, 2.7 microns	Extreme in bands at 4.3, 6.0, 15.0 microns	Extreme in bands at 25 through 1,000 bands	Low except for some specific bands	Very slight
Data type	Analog signal pulse height analysis			Photomul-tiplier sig-nals	Image Orthicons & film	Photographic film (imaging) spectral series			Image & tape	Radio-meter out-put (tape)	Radio-meter out-put (tape)	Radiometer response (tape)	Analog signal, can be converted to imaging system
Storage format	Films or digital tape (nonimaging)			Film or tape (possibly imaging)	Film or tape	Films - which may br read by a flying spot scanner - converted into digital tape			Films after an imaging system	Tape-ana-log signal	Tape-ana-log signal	Oscilloscope or tape & film for storage	Oscillograph or tape & film for storage
Effective source depth of information in re-ceived signal	"Pene-trates 20 cm of rocks"	One to several microns		Angstroms to milli-meters	Angstroms to milli-meters	Only angstroms deep in solids - to tens of meters deep in water; aboorption coefficients best known of any part of the Em spectrum			Source depths are measured in microns (even for water). Some absorption, coefficients are known-attenuation is complete within 10 to 100 microns.			Uncertain; several centimeters only	Sandy- Soil (in.) / Loam (in.) K-Band --- / --- X-Band 2 / 1 C-Band 8 / 2 P-Band 315 / 10
Phenomena detected	Atomic transitions and inner electron shifts			Outer elec-tron shifts	Outer elec-tron shifts	Reflection of visible light; increased contrast; increas-ed haze penetration			Reflectance of solar in-frared-principally surface effects, roughness	Thermal motion of atoms modified by the vibra-tions of the molecules & crystal lattices.		Molecular rotations in atmosphere, surface & subsurface. Emittance is a function of reflec-tance, ab-sorption & temperature	Backscattering by the sur-face; dielectric constant, roughness, depolarization are significant parameters.
Analytical end results	Elemental analysis: Total γ - ray flux; spectral γ-ray flux; ^{40}KU. Th series radionuclides; geo-chemical data.			Elemental analysis: electron transition from ground state	Valence & oxidation, elemental analysis, some gas analysis	Reflectance values (either polychromatic-broad band-pass, or spectrally filtered-narrow bandpass). Struc-tural style, textural data, morphologic data			Molecular compositions. Thin gas layers emit as narrow sharp lines. Solid show broad structureless bands. Textural data, and perhaps some thermal results			Thermal, textural, and com-positional data; emit-tance or apparent tempera-tures dif-ferences; geological value un-defined.	Surface roughness; con ductivity and delectric constant; structural style; textural data

TABLE IV

(Continued)

							Visible			Infrared				
		γ-rays	Hard X-rays	Soft X-rays	Vacuum Ultraviolet	Near UV	Photog UV	Spectral Vis	Photog IR	Near Infrared	Med. Infrared	Far Infrared	Microwave	Radar
Geologic mapping applications	Earth orbit	No application downward			No application downward		Extrapolations from known areas will explain differences in reflectivity (shape, size, and associations form the basis of photointerpretation).			Markedly attenuated by atmosphere; otherwise one obtains reflectivity of solar radiation, roughness criteria	Only some windows available (8-12 μ) restricts compositions of surficial materials which may be determined.	No application downward	Relatively unexplored. Rock diagnosis by difference in apparent temperature-function both of emittance & true temperature difference	Same as for lunar orbit-little or no atmosphere effect.
	Lunar orbit	γ-Ray distribution of nuclear species U, Th, K^{40} distinguishing surface rock types by K^{40} content. Al^{26} Be^{10} may be present if surface has been sputtered. X-ray analysis of a few stable elements			Elemental analysis of surface rocks and dust layers		Photography (visible) cannot give a unique answer to fine-scale physical or chemical compositions unless calibrated areas are used. One can only infer compositions of layers and distributions of shape larger than the limit of resolution.			Sensitivity to solar reflectivities increased because of absence of atmosphere	Surface rock & dust, chemical composition or physical characteristics, apparent temperature range, and compositions not limited by atmosphere	Region unexplored for geoscience value	Near surface and subsurface temperature gradient.	Roughness criteria, perhaps layering in depth (10-1,000 cm), composition, dielectric constant, & particle size differences.
Contribution to major Lunar problems		← Compositions of lunar surface, content of radioactive elements, survey of lunar resources, crater formation; e.g., volcanic action vs. meteorite impacts, extent of lunar rays, dust layer depth, brightness vs. number of secondary corpuscular craters, density of small craters. →								← Thermal state of lunar surface, thermal mapping data, nature of dark spots, e.g., lava lakebeds. →				
Important terrestrial applications		No data possible			← Study of water resources; relationship to industrial growth and general availability; other predictions in hydrology, e.g., precipitation, streamflow, and ground-water levels →		← Data returns for study of metropolitan growth → ← Study of air pollution → ← Optimum use of habitable land → ← Data system for agricultural resource surveys and crop predictions → ← Space derived earth data for use in industrial regions and underdeveloped countries → ← Synoptic analysis of sea state, air-sea interaction, etc. → ← Weather forecasting and air-mass flows →							

[1] Compiled by Peter Badgley (National Aeronautics and Space Administration) and R. J. P. Lyon (Ames Research Laboratory).

TABLE V

POTENTIAL APPLICATIONS OF EARTH RESOURCES DATA GATHERING SYSTEMS

	AGRICULTURAL						GEOGRAPHIC							GEOLOGIC						HYDROLOGIC						OCEANOGRAPHIC					
	VEGETATION DENSITY	GRASS-BRUSH-TIMBERLAND INTERFACES	PLANT SPECIES AND VIGOR	SOIL SERIES, TEMP, & MOISTURE	IRRIGATION WATER	FIRE DETECTION	LAND USE	TRANSPORTATION & LINKAGES	SETTLE & POPULATION MOVEMENTS	RESOURCES UTILIZATION	CLIMATIC CONDITIONS	AIR POLLUTION	TOPO. MAPPING & GEOMORPHOLOGY	COMPOSITION	STRUCTURE	STRATIGRAPHY-SEDIMENTATION	MINERAL DEPOSITS	ENGINEERING	CRUSTAL-MANTLE STUDIES	EVAPOTRANSPIRATION	RAIN DISTRIBUTION & INFILTRATION	GROUND WATER DISCHARGE	WATER POLLUTION	SNOW SURVEYING & GLACIOLOGY	EFFLUENTS OF MAJOR RIVERS	THERMAL CONDITIONS	SEA SURFACE ROUGHNESS	SHOALS & COASTAL MAPPING	BIOLOGICAL PHENOMENA	ICE SURVEILLANCE	SUBSURFACE STRUCTURE
METRIC CAMERA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		●		●	●	●	●	●	●	●	●	
PANORAMIC CAMERA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		●	●	●	●	●	●	●	●	●	●	
MULTISPECTRAL TRACKING TELESCOPE	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		●	●	●	●	●		●	●	●	●	
MULTIBAND SYNOPTIC CAMERA *	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
RADAR IMAGER	●	●	●	●	●	●	●	●		●	?		●	●	●	●	●	●			●	●		●			●	●	●	●	
RADAR ALTIMETER/SCATTEROMETER							?	?					●					●						●			●			●	
WIDE RANGE SPECTRAL SCANNER **	●	●	●	●		●	●					●		●	●	●	●	●	●	●	●	●	●	●	●	●		●	●	●	
IR RADIOMETER/SPECTROMETER	●		●	●		●	●	●						●		●	●	●	●	●		●	●	●	●	●		●	●	●	
MICROWAVE IMAGER				●	●		●	?							●	●			●		●	●	●	●	●	●	●			●	
MICROWAVE RADIOMETER	●	●		●	●	●	?	?						●					●	●	●		●	●	●	●	●	●		●	
LASER ALTIMETER, SCATTEROMETER													●		●	●	●	●	●					●	●		●	●		●	
MAGNETOMETER																	●		●												●
ABSORPTION SPECTROSCOPY												●		●			●	●											●		
RADIO FREQUENCY REFLECTIVITY																●								●						●	
VIEWFINDER ***						●		●		●	●	●			●	●			●		●		●	●	●			●	●	●	
ULTRAVIOLET SPECTROMETER-IMAGER														●	●	●	●														

* Defined for both film return and telemetry modes.

** 0.32 - 14.0 microns.

*** This instrument augments an astronault's vision with optical power and directional data. The astronaut can utilize the viewfinder by itself or in conjunction with other directional type sensors.

	WEEK ENDING																							
	MAY		JUNE				JULY				AUGUST				SEPT.				OCT.					
	21	28	4	11	18	25	2	9	16	23	30	6	13	20	27	3	10	17	24	1	8	15	22	29
BOSTON EDISON		X																						
AAP PRINCIPAL INVESTIGATORS			X							X														
CORNELL UNIVERSITY				X																				
CONRAC				X																				
N.Y. CITY DEPT. AIR POLLUTION CONT.				X											X						X			
U.S. ARMY-FORT BELVOIR						X																		
UNIV. OF NEW MEXICO											X													
CONN. GENERAL LIFE INS.												X												
NEW YORK UNIVERSITY															X									
KOLLSMAN INSTRUMENT															X									
WEYERHAEUSER																X								
KENNECOTT COPPER											X					X								
IBM																X			X					
AMPEX																X								
NASA/AMES RESEARCH																X*								
SHELL OIL																			X					
SHELL DEVELOPMENT																				X				
BETHLEHEM STEEL																						X		
U.S. STEEL																							X	

*requested subsequent return visit

Figure 2 User Presentation Summary

it became necessary to educate the selected representatives of industry in a manner to create a positive image for the pursuance of the AAP experiments data. Of the industrial representatives contacted, only 28 percent were aware of the scope of the experiments or of when the program is slated to get underway. Further, only 17 percent of those contacted were aware of the AAP applications relevant to their own scope of research, only 22 percent were aware of the results obtainable from combining AAP sensor data, and none were aware that data from the AAP would be made publicly available.

The presentations which were delivered were very well received and stimulated much interest. All but one of the companies have since established working groups within their respective organizations to examine the applicability of AAP data to their own product/research effort and have appointed liaison personnel to work with NASA in developing the AAPA. A further indication of the interest stimulated is the fact that these liaison personnel are also, by and large, top management personnel and that their appointments are for an indefinite period. A list of the liaison personnel is provided in Table VI along with a list of those initially contacted as corporate management representatives.

During the user visits, a question and answer session was conducted following the presentation so that some immediate feedback was derived. In every case, the verbal response has indicated a need for more information from NASA about the AAP experiments. Specifically, the information required by industries includes a list of AAP experiments, the types and number of sensors to be employed, the signature characteristics of the sensors, the geographical locations to be scanned by the sensors, and points of contact within NASA concerning experiments.

In summary, most of the potential users contacted are eager for data and already have equipment capable of handling it. The eagerness has demonstrated itself in an expressed desire by some companies to design and submit their own experiments and to participate in a pilot Apollo Applications Program Archives System. However, it is recommeded that NASA release more information to these potential users concerning the details of currently approved experiments. Because of industry's general unawareness of the AAP, the user survey under this contract was forced into the position of first becoming a user Education Program. This proved highly successful and has provided information for a first estimate of data requirements. However, a continuation of the survey is recommended if NASA expects to work with industry toward uses of experimental data derived from AAP. The highly organized method of operation derived during this contract has provided the key elements (i.e., initial contact point, prebriefing letter, presentation, continued liaison) for an operating necleus which can be expanded. In the event that it is impossible to continue the liaison established during this study, action should be taken to dissolve this relationship to prevent loss of confidence in the ability of NASA to deliver required information to the potential user.

TABLE VI
INITIAL CONTACTS AND LIAISON PERSONNEL (Sheet 1 of 2)

COMPANY	INITIAL CONTACT	PROGRAM LIAISON
Boston Edison	Mr. S.J. Sweeney, Director of Operations	Mr. S.J. Sweeney/Mr. B.H. Weiner
Cornell University	Professor A.J. McNair	Professor A.J. McNair
Conrac	Mr. Tom Hewitt	Mr. Tom Harriman, Corporate VP
N.Y.C. Department of Air Pollution Control	Dr. E.F. Ferrand, Director of Technical Services	Dr. E.F. Ferrand
Fort Belvoir	Dr. Colvocoresses	Mr. Orr, Ft. Belvoir, Va. Liaison
University of New Mexico	Mr. Wm. A. Schinick, T.U. Director	Mr. Thomas Lyons
Connecticut General Life	Mr. George W. Young, Sr. V.P.	Mr. H.Z. Lebed, Sec. Actuary's Dept.
Kollsman Instrument Co.	Mr. Jack A. Anderson, V.P. Marketing	Mr. J.R. Downing, Chief Scientist
Weyerhaeuser Co.	Mr. John L. Aram, V. P. Trans. and Materials	Mr. A.S. Gregory, Director of Research
Kennecott Copper Corp.	Mr. Paul Eimon, Sr. Research Geologist	Ralph C. Holmer, Director Exploration Services
IBM	Mr. Ira Lohman, Mgr. Systems Development Lab	Mr. B.W. Wyatt, Mgr. Special Products Market

TABLE VI

INITIAL CONTACTS AND LIAISON PERSONNEL (Sheet 2 of 2)

COMPANY	INITIAL CONTACT	PROGRAM LIAISON
Ampex	Mr. Robert Owen, V.P.	Mr. D.H. Bergis, Planning Mgr.
NASA/ARC	Mr. Edgar Van Vlech	Mr. Harry Hornby
IBM	Dr. Emanual Piore, V.P. and Chief Scientist	Mr. A. Adelman, Federal Systems Division
Shell Oil Company	Mr. E.F. Loveland, Mgr. Public Relations	Dr. Noyes Smith, VP Shell Development
Shell Development Company	Mr. T. Barron, President	Dr. Noyes Smith
Bethlehem Steel Corporation	Dr. D.M. Frazer, Chief Geologist Mining Department	Dr. D.M. Frazer
United States Steel Corp.	Mr. Arthur V. Wiebel, Exec. VP	
New York University	Prof. Philippe Lautrel, EE Dept.	Prof. Philippe Lautrel

3.0 PERFORMANCE SPECIFICATIONS

3.1 General

In order to arrive at the performance specifications for the AAP Experiments Data Archives (AAPA), first it is necessary to decide upon the overall configuration of the basic operational system. Before this can be accomplished, certain governing and influencing parameters must be obtained and evaluated.

The parameters which are discussed in this section may be grouped under three broad categories: input, output, and system environment. Input consists of the experimental data comprising the archives data base: its total volume, quantitative distribution among the various scientific/technical areas involved, rate of accrual, and various formats and media. Output encompasses the total data distributed to the user community. Here is where the feedback concerning the user requirements is brought to bear: the data types, formats, media, and volume which are to be provided will represent a joint consideration of user requirements and the archive's cost effective implementation scheme. System environment parameters define the real world of existing policies, funding, intra- and interagency relations within which the archive must function.

In this section, input considerations are further broken down and enumerated under the headings of Experiment Classification, Data Storage Considerations and Estimate of Data Base. Output is discussed in the paragraph entitled Data Usage Considerations. System environment is essentially an overview which prevails throughout this and subsequent sections.

3.2 Experiment Classification

The Office of Manned Space Flight (OMSF) documentation lists experiments currently included in the NASA Manned Space Flight Experiment Program. The total number of experiments listed (192 at the time of this writing) are classified primarily on the basis of the office sponsoring the experiment (OMSF, OSSA, DOD, etc.), each of which encompasses a number of scientific and technical areas. This classification scheme is presented in Table VII.

It is recognized that the AAPA must preserve a certain measure of continuity relative to the original sponsoring office. DOD, for instance, may decide to impose certain data handling requirements upon its experiments data that differ from other agency requirements. Or, contiguous storage of all experiments sponsored by a particular office may be politically or operationally desirable. These and other considerations, based on the experiment groupings of Table VII must be carefully considered when structuring the archive characteristics.

TABLE VII
EXPERIMENT CLASSIFICATION SCHEME EMPLOYED BY OMSF EXECUTIVE SUMMARY

	Sponsoring Office
DOD	DOD
Space Applications	
Electronic Controls	
Space Vehicle Technology	
Navigation and Traffic	
Particles and Fields	
Engineering Activity	
Biomedical and Behavioral	
Communication	
Scientific	OSSA
Astronomy	
Environment Biology	
Natural Resources	
Meteorology	
Physical Biology	
Particles and Fields	
Solar Physics	
Planetary Atmosphere	
Interplanetary Dust	
Ionospheric Physics	
Planetology	
Communications	
Technological	OART
Communications	
Navigation and Traffic	
Electronic Controls	
Biotechnology	
Space Vehicle Technology	
Biomedical and Behavioral	
Behavioral Technology	
Interplanetary Dust	
Engineering Activity	

TABLE VII (Continued)
EXPERIMENT CLASSIFICATION SCHEME EMPLOYED
BY OMSF EXECUTIVE SUMMARY

	Sponsoring Office
Medical	OMSF
Biomedical and Behavioral	
Engineering	OMSF
Engineering Activity	
Communications	
Navigation and Traffic	
Solar Physics	
Space Vehicle Technology	
Lunar and Planetary SC	

Nevertheless, it can also be seen that some revision of the above classification scheme must be effected in order to better evaluate two additional items of paramount importance, namely: 1) archive organization on this basis of scientific and technical fields of interest and 2) archive organization on the basis of user requirements.

The twenty different scientific and technical areas listed in Table VII offer a sound base for the type of reorganization called for in item 1), except that twenty areas are somewhat bulky to handle effectively. It is advantageous, therefore, to regroup these areas into a smaller number of more general categories. Accomplishing this in light of the criteria of item 2), the various scientific and technical areas have been sequestered into seven major categories which are roughly tantamount to and equatable with the regrouped S.I.C. categories of Section 2.0. These seven categories and their included scientific and technical areas are listed in Table VIII. The degree to which these topical categories relate to the regrouped S.I.C. categories is further discussed in Section 3.5.

Tables VII and VIII provide a working base from which more detailed qualitative and quantitative investigations of the data base can be made. These investigations are accomplished in the ensuing paragraphs.

3.3 Data Storage Considerations

Even a cursory glance through the Executive Summary reveals the wide expanse of scientific and technological endeavor encompassed by the currently planned experiments. This leads to numerous questions concerning the data base which must be resolved in order to establish a functioning and effective archive. However, as will subsequently be seen, not all of these questions are completely resolvable at this time, due to the tentative definition of substantial areas of Apollo Applications. Consequently, the goal of this study has been to investigate all areas within the bounds of presently available information, resolve those areas which appear sufficiently well defined, and leave posed, but unresolved, those questions which appear to be inadequately defined at this time. In doing so, it is hoped that presently nebulous areas may not be deceptively blanketed by gross engineering estimates. Instead, by properly underscoring the less defined areas, it becomes possible to design an open-ended system with the major points of flexibility concentrated in the unresolved areas.

3.3.1 Types of Experiments to be Housed. - Not all of the experiments data collected during the course of the AAP will necessarily be deposited in the Experiments Data Archive. There are two predominant factors which will influence the archive allocation of this data; namely, the relationship of the AAPA with existing archives, and the consideration of user requirements.

3.3.1.1 Relationship with Existing Archives. - There are several existing archives within the structure of the Federal Government which are dedicated to the storage

TABLE VIII
CATEGORIZATION OF AAP SCIENTIFIC AND TECHNICAL AREAS

Assigned Category	Inclusive Scientific and Technical Areas
Space Science	Astronomy Planetary Atmosphere Interplanetary Dust Planetology
Physical Sciences	Solar Physics Particles and Fields Ionospheric Physics
Space Engineering	Engineering Activities Space Vehicle Technology Electronic Controls Space Applications Navigation and Traffic Lunar and Planetary SC
Life Sciences	Biomedical and Behavioral Environment Biology Physical Biology Biotechnology
Communications	Communications
Meteorology	Meteorology
Earth Resources	Natural Resources

of data or information endemic to certain specialized areas of science or technology. Any of these archives are possible candidates for the storage of those portions of the AAP experiments data falling within the bounds of their individual specialties.

One case of particular interest during the course of this study was the National Space Science Data Center (NSSDC) located at the Goddard Space Flight Center in Greenbelt, Maryland. This center was established by NASA to further the widest practicable use of reduced data obtained from space-science investigations and to provide investigators with an active repository for such data. As such, it is responsible for the active collection, organization, storage, announcement, retrieval, dissemination, and exchange of data received from satellite experiments, sounding-rocket probes, and high altitude aeronautical and balloon investigations. In addition, the Data Center collects correlative data, such as magnetograms and ionograms, from ground-based observatories and stations for NASA investigators and for onsite use at NSSDC in the analysis and evaluation of space science experiment results. NSSDC is, therefore, the logical repository for all space science data resulting from AAP experiments. This organization has selected forty-two AAP experiments which are within NSSDC's field of interest, and has listed twenty-five others for which there is presently insufficient information to base a decision*. Up to thirty-five percent, then, of the presently scheduled AAP experiments should be housed at NSSDC.

In addition, it was recommended that this study examine the National Weather Records Center in Asheville, North Carolina; the National Oceanographic Data Center (Navy) in Washington, D. C., and the National Aeronomy and Space Data Center (ESSA) in Boulder, Colorado, as possible repositories for other selected experiments. As brought out earlier by ERC**, both the National Library of Medicine and the National Library of Agriculture would also be candidates for selected AAP experiments results. Basically, the option which exists is that of either incorporating certain of these repositories within an AAP archives system, and thus topically decentralizing the achives itself, or treating these other repositories as AAP Data Users, and establishing a more centralized archives structure for AAP Experiments Data. A more detailed evaluation of the various ramifications involved in this trade-off is presented in Sections 4.0 and 5.0 of this report.

It is advantageous at this point, however, to consider only a single division between experiments on the basis of topic, namely: space science or non-space science oriented. The breakdown is shown in Table IX. In some cases, an entire category has been relegated to NSSDC for the sake of continuity. In other areas, some duplication of storage

---

*NASA letter, dated August 28, 1967 from James Vette, NSSDC, to Maurice Moroney, ERC.
**The ERC/SL Presentation at the AAP Data Flow Planning Review at the Marshall Space Flight Center, March 28, 1967.

TABLE IX

TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
I. SPACE SCIENCE			
A. Astronomy			
*S001	Zodiacal Light Photography	100	
*S013	UV Astronomical Camera		
*S017	X-Ray Astronomy		
*S019	UV Stellar Astonomy		
*S027	Galactic X-Ray Mapping		
*S028	Dim Light Photography		
*S030	Dim Sky Photography/or Thicon		
*S069	X-Ray Astronomy (B)		
B. Planetary Atmosphere			
S011	Airglow Horizon Photography	50	100
+S021	Airglow Horizon Photography		
S025	IR Temperature Soundings		
*S063	UV Airglow Horizon Photography		
C. Interplanetary Dust			
*S020	UV X-Ray Solar Photography	100	
*S035	Medium Energy Solar Wind		
*S038	Low Energy Solar Wind		
*S064	UV Photography of Dust Clouds		
*S070	UV/X-Ray Solar Photography		
*T015	Meteoroid Composition		
*T016	Meteoroid Entry Observation		
*T017	Meteoroid Impact + Erosion		
*T021	Meteoroid Velocity		
D. Planetology			
*S031	Lunar Passive Seismology		
*S032	Lunar Gravity		
*S033	Lunar Active Seismology		

*Desired by NSSDC
+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
I. SPACE SCIENCE			
D. Planetology			
*S034	Lunar Tri-Axis Magnetometer		
*S036	Suprathermal ION Detection		
*S037	Lunar Heat Flow	100	
*S058	Cold Cathode Ionization Gauge		
*S059	Lunar Geology Investigation		
II. PHYSICAL SCIENCES			
A. Solar Physics			
M454	Shadow Shield		
*S010	Agena Micrometeorite Collect		
S012	Micrometeorite Collection		
*S018	Micrometerorite Collection		
*S029	Libration Regions Photographs		
*S052	White Light Coronagraph	100	
*S053	UV Coronal Spectographs		
*S054	X-Ray Spectrographic Telescope		
*S055	UV Spectrometers		
*S056	Dual X-Ray Telescopes		
+S062	Solar X-Ray Flux Variation		
+S067	Gamma Ray and X-Ray Spectroscopy		
B. Particles and Fields			
+S009	Nuclear Emulsion		
*S016	Trapped Particles Asymmetry		
*S022	Low Z Cosmic Rays		
*S023	High Z Cosmic Rays	100	
+S024	Nuclear Emulsion		

*Desired by NSSDC
+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
II. PHYSICAL SCIENCES			
B. Particles and Fields			
*S051	Daytime Sodium Cloud		
D008	Radiation in Spacecraft		
C. Ionospheric Physics			
+S026	ION Wake Measurement	100	
III. SPACE ENGINEERING			
A. Engineering Activities			
M401	Mapping and Survey System		
M402	Orbital Workshop		
M403	Electrostatic Charge		
*M404	Proton/Electron		
*M405	Tri-Axis Magnetometer	14	86
*M406	Optical Communication		
*M407	Lunar UV Spectral Reflect		
*M408	Beta Spectrometer		
*M409	Bremsstrahlung Spectrometer		
M410	Color Patch Photography		
M411	2-Color Earth S Limb Photos		
M412	Landmark Contrast Meas		
M413	Subcritical Cryo Storage		
M415	Thermal Control Coatings		
M416	Propell and Mass Determination		
M417	Liquid Interface Stability		
M418	Boiling Heat Transfer		
M419	Cryogenic Propellant Transfer		
M420	Propell and Storage System		
M422	DC Motor + Gear Lubrication		

*Desired by NSSDC
+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
III. SPACE ENGINEERING (Continued)			
A. Engineering Activities (Continued)			
M423	Hydrostatic Gas Bearing		
M426	Condensing Heat Transfer		
M427	Strapdown Platform		
M432	Large Space Structures		
M433	Satellite Recovery		
+M436	Artifical Gravity		
M466	Space Suit Evaluation		
M469	ST-124 Removal		
M475	Water + Waste Management		
M484	Orbital Workshop Artificial G		
M486	Astronaut Eva Equipment		
M487	Habitability/Crew Quarters		
M492	Tube Joining in Space		
M493	Electron Beam Welding		
M495	Lunar Surface Engr Properties		
M496	Liquid Drop Dynamics		
M497	Fluid Density Gradient		
M508	Astronaut EVA Hardware Evaluation		
M509	Astronaut Maneuvering Evaluation		
1002	Moderate Depth Drill		
1004	Lunar Surveying System		
1010	Variable Gravity Support		
T019	Pegasus Panel Retrieval		
T020	Jet-Shoes		
D012	Astronaut Maneuvering Unit		
D016	Power Tool Evaluation		
D017	Carbon Dioxide Reduction		
D018	Integrated Maintenance		
D019	Suit Donning + Sleep Sta Eval.		
D020	Alternate Restraints Eval.		

+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
III. SPACE ENGINEERING (Continued)			
B. Space Vehicle Technology			
M479	Zero Gravity Flammability		
M483	Radar Attitude Sensing System		
M488	High Pressure Gas Expulsion		
M489	Heat Exchanger		
T005	Fusible Material Radiator		100
T022	Heat Pipe		
T023	Surface Absorbed Materials		
D003	Mass Determination		
D021	Expandable Airlock Technology		
D022	Expandable Reentry Structures		
C. Electronic Controls			
T003	In-Flight Nephelometer		
T011	Reentry Communications		100
T013	Crew-Vehicle Disturbance		
D002	Nearby Object Photography		
D. Space Application			
D001	Basic Object Photography		
D006	Surface Photography		100
D007	Space Object Radiometry		
D015	Night Image Intensification		
E. Navigation and Traffic			
+M439	Star Horizon Automatic Tracking		
M411	Noncooperative Tracking Radar		
T002	Manual Navigation Sightings		
T014	Orbital Horizon Definition	25	75
+D004	Celestial Radiometry		
D005	Star Occultation Navigation		

+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION(%) NSSDC	AAPA
III.	SPACE ENGINEERING		
E.	Navigation and Traffic		
D009	Simple Navigation		
D010	ION-Sensing Attitude Control		
F.	Lunar and Planetary SC		
1001	LSSM		100
1009	Nuclear Emulsion		
IV.	LIFE SCIENCES		
A.	Biomedical and Behavioral		
M001	Cardiovascular Conditioning		
M003	In-Flight Exerciser		
M004	In-Flight Phonocardiogram		
M005	Bioassays Body Fluids		
M006	Bone Demineralization		
M007	Calcium Balance Study		
M008	In-Flight Sleep Analysis		
M009	Human Otolith Function		
M011	Cytogenetic Blood Studies		
M012	Exercise Ergometer		
M017	Thoracic Blood Flow		100
M018	Vectorcardiogram		
M019	Metabolic Rate Measurement		
M020	Pulmonary Function		
M021	Semicircular Canal Function		
M022	Red Blood Cell Survival		
M023	Lower Body Negative Pressure		
M048	Anti G Elastic Garment		
M049	Analysis CO + CO_2		
M050	Metabolic Activity		
M051	Cardiovascular Function Assess		
M052	Bone and Muscle Changes		

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
IV.	**LIFE SCIENCES**		
A.	Biomedical and Behavioral (Continued)		
M053	Human Vestibular Function		
+M054	Neurological Study (EEG)		
M055	Time and Motion		
T006	Vision Test Equip Eval		
T007	Human Transfer Functions		
D013	Astronaut Visibility		
B.	Environment Biology		
S002	Sea Urchin Egg Growth		
S004	Radiation + Zero G on Blood		100
S061	Potato Respiration		
T009	Primates in Long Term Zero Gravity		
C.	Physical Biology		
S008	Visual Acuity		100
S015	Zero G Single Human Cells		
D.	Biotechnology		
T004	Frog Otolith Function		
T008	Electrolysis Cell		100
T010	On-Board Centrifuge		
V.	**COMMUNICATIONS**		
A.	Communications		
M430	LM Relay		
M440	Wide Band Variable Power Trans		
1003	Emplaced Scientific Station		
S041	Millimeter Wave Propagation		100

+Possibly Desired by NSSDC

TABLE IX
TENTATIVE ARCHIVE ALLOCATION OF ANNOUNCED AAP EXPERIMENTS
(Continued)

EXPERIMENT		ARCHIVE ALLOCATION (%) NSSDC	AAPA
V. COMMUNICATIONS			
A. Communications (Continued)			
T001	Reentry Communications		
T012	Optical Technology		
+D014	UHF-VHF Polarization		
VI. METEOROLOGY			
A. Meteorology			
S006	Synoptic Weather Photography		
+S007	Cloud Top Spectrometer		
*S039	Day-Night Camera System		
S040	Dielectric Tape Camera System		
+S043	IR Temperature Sounding	As Desired	100
+S044	02 + H20 Microwave Radiometer		
*S045	IR Filter Wedge Spectrometer		
S046	Visible RAD Polarization Meas		
S047	Stellar Refract Density Meas		
S048	UHF Sferics Detection		
+S049	IR Interferometer Spectrometer		
+S050	15 Micron Grating Spectrometer		
+S057	Multi-Channel Radiometer		
+S060	Selective Chopper Radiometer		
VII. EARTH RESOURCES			
A. Natural Resources			
S005	Synoptic Terrain Photography		
+S042	Multi-Spectral Photography	As Desired	100
+S065	Multi-Bank Terrain Photography		

*Desired by NSSDC
+Possibly Desired by NSSDC

(overlap) may prove desirable in order to better service both AAPA and NSSDC users. This table affords a cursory insight, by major category, to the particular experiments data which may be expected to be housed at NSSDC or the AAPA, assuming that the AAPA houses all nonspace science data. The considerations for topically decentralizing the body of nonspace science data are enumerated in Section 4.0, while the actual relationships between the AAPA and NSSDC are discussed in Section 5.0.

3.3.1.2 Consideration of User Requirements. - One of the principal concepts underlying the establishment of the AAPA is that "user needs [be] recognized as key considerations because they affect the amount of data stored in the archives and also the frequency with which this stored data is used."* To help ascertain these needs, questionnaires were distributed to the companies and institutions which received briefings from the AAP user requirements team. Although many of the more specific questions could understandably not be answered at the time, some general requirements were expressed and others can be inferred. For instance, some categories of experiments, namely those classified under Earth Resources and Meteorology, are of such a nature that the sensors involved and their method of employment should enable the resultant data to be utilized by a wide variety of users in a manner never intended or conceived by the original experimenter. This will undoubtedly hold true both for other whole categories and for various isolated experiments. On the other hand, because of the high degree of specialization of many other scheduled AAP experiments, it may be predicted with equal certainty that there will be a significant number of experiments which will contain little or no "user appeal". This also has been borne out by discussion with potential users.

Despite their obvious importance, however, consideration of user needs does not alone define the archive. Some means of archival storage must be provided for all AAP experiments data, regardless of the presence or absence of any potential user application. The actual problem definition, therefore can be seen to be twofold in that:

1) All experiments data must be archived.

2) Certain selected experiments data must be provided with sufficient retrieval capability to adequately satisfy user needs.

Although these different aspects of the problem have sufficient commonality to make a single solution feasible, consideration of dual, but related, solutions may be warranted.

*ERC/SL Presentation

If a single homogeneous solution is chosen, the problem of anticipating user requirements is effectively eliminated; all experiments are stored on an equal basis, and the user has complete freedom of access to any which interest him. It necessitates, however, that the entire data base be "active" to the same degree as the most frequently accessed data. This poses a considerable burden on the entire retrieval system, the complexity of which tends to be directly proportional to the size of the data base.

On the other hand, if a dual solution is chosen, whereby the data base is divided into "active" and "inactive" sections, the archival requirements are still satisfied, yet the retrieval burden - since it would operate primarily from the "active" data base - is considerably alleviated. The main problem then, is how to decide which data should be active, and which inactive. Although the basic criterion is obviously "user requirements", the means of implementing this criterion are unfortunately limited, and require a closer examination. For example, it might be argued that the body of experiments to be housed in the "active" AAPA be preselected on the basis of their anticipated user appeal. However, it must be remembered that there are presently not adequate means for establishing in advance the complete usefulness of the experiments data even to the experimenter himself*.

Selection might be postponed, then, until after the data has been recovered and subsequently reduced and analyzed by the experimenter. At that point, the potential usefulness of the data would be better known and a more meaningful set of criteria for assaying its potential user appeal could be established. Nevertheless, despite the soundness of any preestablished criteria, there will always be some potential application of data which will be overlooked by any review board, no matter how competent. This is due to the inherently inventive nature of the type of applications technology towards which Apollo Applications is geared.

Since any advance prediction of user requirements can be only a gross approximation of the actual requirements, it is worthwhile to consider a system which is more closely coordinated with these actual needs. Such a system could operate as follows:

1) Accept, process as necessary, catalog, and actively store all incoming experiments data. **

---

*Ludwig, A. H. GSFC, Advanced Space Information Systems

"It must be kept in mind that the experimenters are frequently not able to predict the detailed characteristics of their expected data accurately enough to permit them to write the final processing programs before launch."

**The topical apportionment of experiments discussed in the previous subsection is intentionally ignored at this point for the sake of simplicity.

2) Maintain each experiment in active storage for some set period of time, and record its usage characteristics.

3) If any experiment fails to meet a predetermined usage count, or subsequently falls below a certain usage rate, purge it by transferring it from active to inactive (or less active) storage. These "data retirement thresholds" may be left variable, and thus additionally function as a control over data base stability.

4) As experiments (or portions thereof) are retired, record this fact in a catalog so that there remains some centrally located record of all experiments which have passed through the system. After some longer period of time, certain sections of the active catalog may themselves be retired.

In summation, it has been concluded that although user requirements must be given due consideration in determining which experiments data are to be stored in the AAPA, the exact requirements cannot be predicted in advance with a full degree of accuracy. Nor is across-the-board active storage of all data the solution; this operates at cross purposes with the effective and economic retrieval of data required to satisfy user needs. The argument then, is strongly in favor of a dynamic system involving the concepts of both active and inactive (or less active) storage, and depending on the retirement of low usage data for preservation of effective retrieval. This is the basis of the system concept outlined in paragraph 3.8 of this section.

3.3.2 Forms of Data to be Stored. - The AAPA will have to deal with many different forms of data in order to house a wide variety of experiment categories and satisfy the demands of an equally diversified field of users. Specifically, it must be determined:

- Which forms will be accepted by the archives
- Which forms will actually be provided archival storage
- Which forms will ultimately be distributed to the user

The underlying considerations which must first be resolved in order to determine the above are twofold, namely: 1) which data forms are normally produced during an experiment, and 2) which forms will be required by the user.

It has been noted that information relating to item 2) is extremely sketchy at this time. This situation may be perpetuated to a certain extent well beyond the first AAP launches. Consequently, it is deemed advisable that - initially at least - users be constrained to utilize either the same data made available to the PI, or that which is normally yielded by PI reduction and analysis. This is considered to be a workable constraint in that: 1) the data provided for the experimenter is normally in a format adaptable to large-scale digital computers*, and 2) the initial users of the AAPA may be expected to be primarily those organizations already having comparable computer capability at their disposal. If it is subsequently determined that additional formats are required to properly serve an expanding user community, this service may be effectively incorporated at the retrieval and output portion of the archive, rather than at the input and storage portion.

This results in a fortunate simplification, namely, that the data forms to be accepted and stored by the AAPA consist principally of those forms normally made available to the PI, or those produced by his subsequent reduction and analysis. In order to more competently evaluate the AAPA data storage requirements, a review of both the types of experiments data and their various format versions is required.

3.3.2.1 Data Types to be Stored. - In spite of the increasing complexity of the sensors employed in space experiments, coupled with the greater involvement of onboard computer processing, the variety of methods actually required to record the resultant experiments data appears relatively fixed. Analog, sampled analog, bi-level, digital, and scanned image (facsimile or digital representation) data is normally transmitted via PCM telemetry to one of the Manned Space Flight Network (MSFN) Remote Sites. The signal is then recorded, along with ground timing and other pertinent information, on seven track analog magnetic tape. After telemetry post-processing, "cleaned up" digital versions are forwarded to the experimenters. Alternately, any of the above data types can also be stored on magnetic tape aboard the spacecraft for eventual recovery after splashdown. In this case, image data can be recorded photographically, rather than by facsimile techniques. Other data may appear in the form of astronaut notes and physical samples.

Under contract with NASA, Bellcomm, Inc., has compiled a list of experimenter data requirements for the initial five AAP launches involving fifty-five experiments**. This report lists the recovered data as being of the following types:

Electrical

Digital - Expressed in bits

Analog -

Narrow band (0 - 100 Hz) - Sampled for transmission and recorded digitally.

---

*Telemetry Processing for NASA Scientific Satellites, Habib, E. J., Information Processing Division, GSFC.

**Data Generated by Experiments on AAP Flights 1A through 4, Agarwal, R.K., Bellcomm, Inc., Sept. 22, 1967.

Medium band (100 - 3000 Hz) - Recorded as analog, expressed in medium band channel hours.

Wide band (> 3000 Hz) - Recorded as analog, expressed in wide band channel hours.

Non Electrical

Film - 8, 16, 35, 70, and 75 mm, plus "film strips" and "photos"*.

Written notes - Hard copy

Physical Samples - Biological, soil, and space samples (e.g., stardust and green cheese)

For the purposes of subsequent data volume estimates, plus the overall AAPA system design, the above data types are considered to represent the input types requiring storage within the AAPA. The actual system media for effecting this storage is discussed in Section 6.0.

3.3.2.2 Data Versions to be Stored. - Between the time an experiment is conducted on board the spacecraft and the time the PI publishes his final results, the experiment data will undergo an assortment of data handling and processing operations. During the course of these operations, several different versions of the original data will be produced: some of these constitute merely refinements or a rearrangement of the base data, while others involve data reduction and interpretation. A review of some of the aspects of post-flight data flow is therefore required in order to 1) enumerate the major data versions produced, and 2) ascertain which versions are best suited for the AAPA.

The Raw Analog Data tapes recorded at the MSFN sites and containing the telemetered experiment data will be forwarded to the GSFC, where they will be monitored to assure proper network operation. These tapes will then be forwarded to MSFC or MSC for processing. NASA Center processing will consist of the following activities: **

1) Analog-to-digital processing, in which the data will be signal conditioned, converted to digital form and recorded on digital magnetic tapes.

---

*The conventional standards for photographic film, as established by MSC, are: 8, 16, 35, and 70 mm, 5 in., and 9 1/2 in.

**AAP Experiments Data Flow Plan, Bellcomm, Inc.

2) Digital computer processing, in which the data will be processed to flag erroneous data readings, to verify the correct time when necessary and to sort the data readings and time for presentation to the appropriate experimenters.

3) Data plotting, in which data to be presented to experimenters in graphical form will be plotted.

4) Support data processing, in which supporting information such as orbit/attitude data and command data will be prepared.

5) Quality control, in which all processed data will be inspected and certified prior to being released to experimenters.

6) Production control, in which the overall scheduling and control of the facility will be achieved.

Figure 3 shows this overall processing sequence as it relates to the NASA Centers. The decommutated (sorted) digital data tapes which are ultimately forwarded to the experimenter are referred to as Experimenter Tapes.

The data retained on board the spacecraft, rather than being transmitted to the MSFN sites, will be shipped with the spacecraft to the MSC. Physical samples will be handled in accordance with instructions from the PI; film will either be delivered intact to the PI or processed at MSC; and on-board tapes will be processed as necessary for release to the experimenter. Together with the Experimenter Tapes these data constitute the Raw Data which will normally be forwarded to the PI.

This Raw Data then passes under control of the PI, who is faced with the task of reducing the bulk data into forms from which meaningful conclusions about the phenomena being investigated can be reached. Reduced Data is defined as data which has been configured in the manner most amenable to final analysis, but has not been subjected to any types of reduction which might alter or destroy any of the fundamental information content. It therefore represents the most refined level of data prior to the interpretation stage. Although the actual point at which data ceases to be "reduced" and commences being "analyzed" may elude precise definition, data reduction commonly incorporates such operations as reformatting, sorting, merging, accumulation, statistical analysis, and mathematical manipulation of the raw data. It also includes some provision for outputting the summarized data in a readable form, such as line printer tabulations, x-y plots, motion pictures, etc. The volume of the data which must be handled, and the complexity of the programs for handling them, vary depending upon the intended analysis technique. These techniques can generally be sorted into four major categories: *

---

*Advanced Space Information Systems, Ludwig, G. H., GSFC, April, 1967.

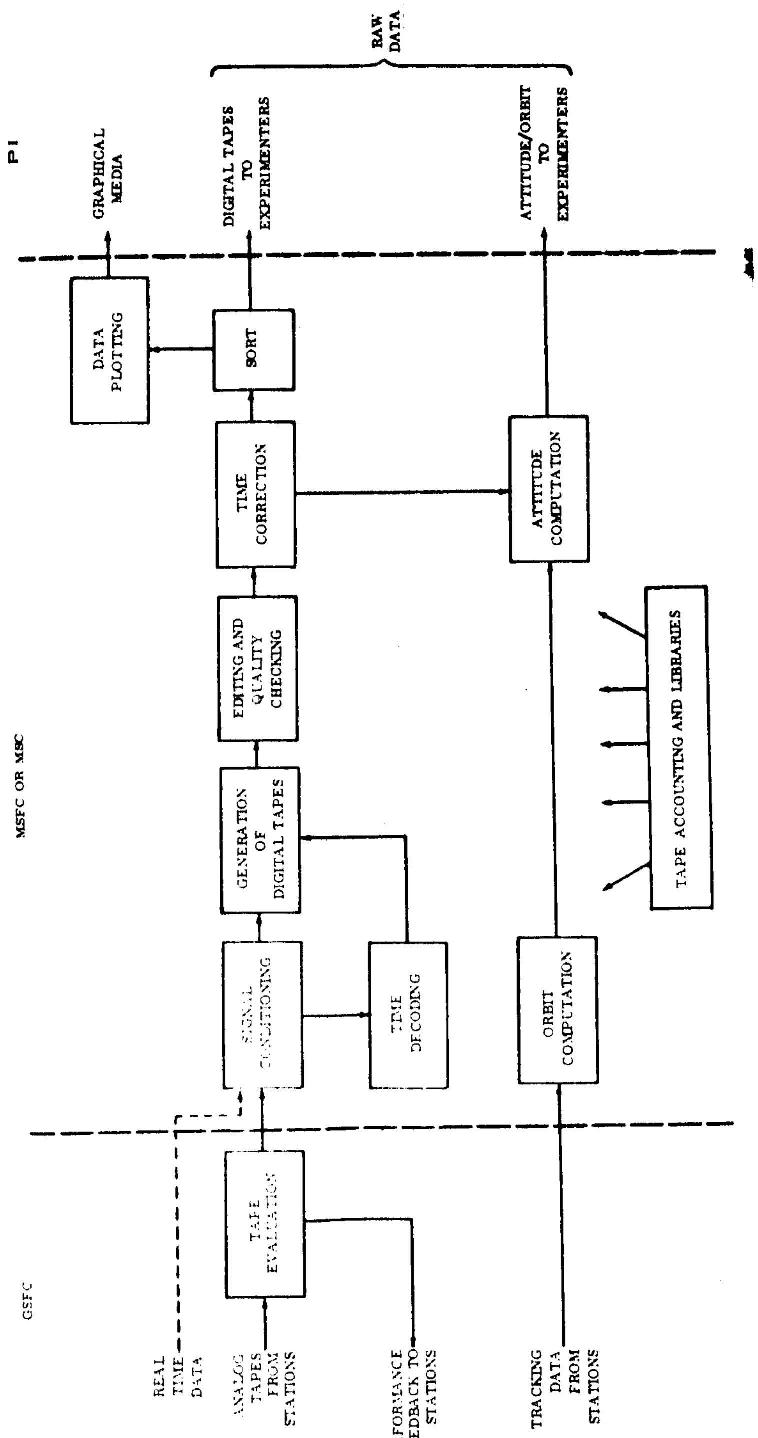


Figure 3 NASA Center Telemetry Data Processing for Experiments Data

1) Data Scanning - It is frequently necessary to scan all or a major portion of the data rapidly and with relatively simple processing to select the most interesting portions for detailed analysis. Dynamic visual displays appear most promising for this operation.

2) Mapping - Many experiments require the selection of data from specific and predictable regions or periods of time to provide a mapping in either space or time. In many of these cases only a portion of the data must be processed, but the processing may be relatively complex.

3) Analysis of Selected Events - Detailed analysis may be performed on selected special occurrences, such as penetration of the earth's magnetospheric boundary surface, or solar storms.

4) Full Analysis - Certain experiments may require full analysis of all data. This is especially true for very slowly occurring phenomena such as micrometeorites or heavy cosmic rays, where all events must be analyzed to obtain statistical significance.

The results of these analyses - the Final Analyzed Data - is that data which the PI designates as best containing the scientific results of his experiments. The end product - namely the experiment results - are ultimately published in the scientific and technical journals serving the PI's particular field of endeavor.

The post-flight experiments data, then, can generally be considered as being processed into four major versions: Raw Analog Data, Raw Data (digital), Reduced Data, and Final Analyzed Data. Obviously, it will not be necessary on the basis of either archival requirements or user considerations to house all of these versions. In the case of the Raw Analog Data, the NASA centers performing the post flight processing will retain the analog station tapes for six months, and then erase and reuse the tapes. *

In assaying the remaining three versions, it is worthwhile to consider the precedent established by NSSDC. To quote from their literature:

> "Primarily, NSSDC acquires reduced data records which have been prepared from the original data by the introduction of such factors as calibration and correlation. These data normally present the value of the quantity measured as a function of time, position, and other appropriate parameters. In addition, an attempt is made to assure later investigators independent use of the data by providing corollary information such as:

---

*AAP Experiments Data Flow, Bellcomm, Inc., Pauley, R. J.

- Description of the instrument or measuring device for collecting data.

- Discussion of any important or unusual developments occurring during the course of the experiment which affected the data.

- Specific methods employed by the principal investigator in performing his data reduction and data analysis.

- References to the scientific results of the analyzed data and all other pertinent bibliographic materials.

In addition, NSSDC obtains those final analyzed reduced data which the principal investigator designates as best containing the scientific results of his experiment. Paramount in this acquisition is the preservation of the integrity of the investigation. Consequently, the form of the reduced data accepted by the Data Center is determined by the needs of the experimenters."

Experience has shown that at NSSDC the Reduced and Final Analyzed Data were the most useful versions of data from the user standpoint. It was pointed out that the Raw Data, or Experimenter Tapes, were normally meaningful only to the experimenter, and that only after the type of reduction outlined above could the data be readily interpreted by other investigators. Of course it must be remembered that in this case only space science experiments data is being considered.

Aside from user requirements, there is another factor which may have to be considered in deciding whether to store Raw Data in addition to the Reduced and Final Analyzed Data in the AAPA. NSSDC has the advantage of being located at Greenbelt, Maryland, along with the GSFC Central Processing Facility (Building 23) and its tape archival facility (Building 16). Under the present policy, this tape storage includes the original analog tapes, buffer tapes, edit tapes, and final experimenter tapes. From the archival standpoint, then, there is no need for NSSDC to store other than the Reduced and Final Analyzed Data versions.

This may not be the case with the AAPA. The NASA Center responsible for the majority of the experiments and subsystems on a module, generally the Center responsible for the module telemetry system, will process all the post-flight data from the module. Consequently, nearly all AAP postflight data handling will be performed either at MSC or MSFC.* Therefore, determination of which data are to be stored in the AAPA depends to a certain extent on the availability of other archival facilities. If there are none available, the AAPA might be configured to store all experiments data except the analog station tapes, which are to be erased and reused. If other archive facilities are available, it might be feasible to store only - the Reduced and Final Analyzed Data - much in the same manner as NSSDC.

Since this item represents a major consideration in deciding the location, decentralization, and organizational structure of the archive, its more detailed evaluation is included in Sections 4.0 and 5.0. For the purpose of defining the archives data base, however, two versions will be considered - one containing the Raw (digital), Reduced, and Final Analyzed Data, and the other just the Reduced and Final Analyzed Data. It will be seen that the only principal difference is the size of the data base; the structure and activity of the system are not significantly affected.

3.4 Estimate of Digital Data Base

The preceding sections have dealt with the types of experiments and versions of experimental data which appear most appropriate for storage in the AAPA. In order to ascertain the characteristics and size of the data base predicated by these considerations, it is first necessary to regress one step and consider the overall experimental data volume.

The AAP Data Flow Plan Tasks and Assignments lists six basic tasks. Task No. 1 - Estimates of Experimental Data Generation - has been undertaken by Bellcomm, Inc. To date, the estimates released by Bellcomm relate to the digital data produced by the fifty-five experiments scheduled for the first five flights. At present, there is not enough information known about the photographic requirements of these experiments to permit an estimate of image data volume.**

In the above referenced report, low-estimate figures are obtained by tabulating only the desired data from each experiment; i.e., the bits that are collected if the duration of the experiment and the rates at which the experiment sensors are sampled are exactly as requested by the PI. These figures are listed in Table X.

---

*AAP Experiments Data Flow Plan, Bellcomm, Inc.

**"Data Generated by Experiments on AAP Flights 1A through 4," Case 600-2, Agarwol, R. K., Bellcomm, Inc., Sept. 22, 1967.

TABLE X
DESIRED DIGITAL DATA VOLUMES FOR AAP
FLIGHTS 1A THROUGH 4

FLIGHT	DATA RATE (Bits/Sec.)	TOTAL DATA (Bits)
AAP 1A	3K	4×10^8
AAP 1-2	18K	2×10^9
AAP 3-4	19K	6×10^9

The actual amount of data collected from an experiment may considerably exceed this "requested" amount because of the small variety of sampling rates available on existing Apollo and Gemini PCM systems. Using this and other related considerations, the report further provides estimates to show the "best-fit" between desired data and the sampling rates available on the carrier module. These more realistic figures are contained in Table XI.

TABLE XI
ADJUSTED DIGITAL DATA VOLUMES FOR AAP
FLIGHTS 1A THROUGH 4*

FLIGHT	DESIRED DATA RATE (Bits/Sec)	ADJUSTED DATA RATE (Bits/Sec)	ADJUSTED TOTAL (Bits)
AAP 1A	3K	11K	1.5×10^9
AAP 1-2	18K	47K	5.2×10^9
AAP 3-4	19K	40K	1.3×10^{10}

Furthermore, it is possible to obtain high estimate figures by considering the spacecraft communications link as the final limiting factor in the collection of experimental data. This approach was suggested early in this study during discussions between Bellcomm, ERC, and UACSC. ** Since the previously referenced Bellcomm report does not attempt an overall data estimate on this basis, this study group has

---

*Adjusted totals are not contained inthe Bellcomm report. Those presented were calculated in direct proportion to the increased data rates.

**UACSC Trip Report (Contract No. 7), Bellcomm, Inc., June, 1967.

made such an estimate based on communications information contained in a second Bellcomm report. *

Using the data figures obtained by Bellcomm for the first 55 experiments and the three approaches outlined above, it has been necessary to establish some bases for extrapolating these figures for the remainder of the projected AAP experiments, and estimating that portion which would be stored in the AAPA. These bases are discussed in the following paragraphs. A thorough familiarity with both of the previously referenced Bellcomm reports is recommended in order to maintain proper perspective throughout the remainder of this section.

3.4.1 Desired Data Estimate. - Table IX was established as a basis for differentiating between space science and non-space science experiments. This same method can also be employed to differentiate between the data volumes of these two factions. Appendix B has been constructed for this purpose. Although hardly an exacting operation, the extrapolation of these initial data figures through the completion of the AAP can be most realistically made by basing one assumption: that the greatest data volume similarity exists between experiments in the same scientific/technical category. Therefore, the most reasonable estimates may be derived by linearly extrapolating the available figures in each scientific/technical category to obtain category subtotals, and then combining these subtotals to arrive at a total for all experiments listed in the Executive Summary. As accomplished in Appendix B, this represents the desired data estimate for the 192 scheduled** experiments.

3.4.2 "Best-Fit" Data Estimate. - Tables X and XI compare the acquired data rates and volumes for "desired" versus "best-fit" versions over the first five flights. The best-fit version can be seen to exceed the desired version by a factor of approximately 2.4. Applying this factor "across the board" to the desired data figures of Appendix B yields the best-fit data estimate for the experiments scheduled in the Executive Summary. Assuming a total of 400*** experiments over the course of the AAP, these estimates may be further extrapolated (linearly) to give the following figures:

NSSDC: 1.5×10^{11} bits

AAPA: 2.2×10^{11} bits

TOTAL: 3.8×10^{11} bits

*"Apollo Applications Program Experiments Data Flow Plan", Pauly, R. J., Bellcomm, Inc.
**Scheduled at the time of Bellcomm's estimate.
***UACSC Trip Report, Manned Space Center, Houston, Texas; Question No. 4.

3.4.3 Telemetry System Transmission and Acquisition Capabilities. - There are four spacecraft module telemetry systems which transmit real-time data and three which transmit playback data from on-board tape recorders. Since various combinations of these modules will be used on different AAP missions, the data transmission capability on a particular mission will depend on the specific modules assigned. For the solar telemetry mission, where all seven systems will be employed, Bellcomm has calculated that the overall daily data transmission capability will be approximately 2×10^{10} bits.* The volume of data acquired at the MSFN ground sites, however, is dependent on ground station coverage - that portion of the mission during which the spacecraft is in contact with at least one ground site. For 30 percent station coverage, the overall daily acquisition capability, as computed by Bellcomm, is 6.68×10^{9} bits.

It will be noted by comparison with Table X that this theoretical daily limit employing all seven module telemetry systems is greater than the mission total for any of the individual launches, or even the combined, 56-day duration, AAP 3/4 mission. Nevertheless, one of the overall program goals during the course of AAP will be full utilization of available hardware, and this includes communications hardware. Assuming a program total of 25 launches, and further assuming that, after the presently scheduled five launches are completed, this upper limit of data acquisition could be attained, it is possible to consider some upper limit to the data totals. If these 20 missions were to be paired into ten twin cluster missions, each having all seven module telemetry systems and a mission duration of 28 days,** the total experiments data collection over the course of the AAP would be on the order of 2×10^{12} bits. By proportioning this figure to the data totals given in paragraph 3.4.2, this high estimate data total can be apportioned as follows:

NSSDC:	0.8×10^{12}
AAPA:	1.2×10^{12}
TOTAL:	2×10^{12}

Comparing these totals to those of paragraph 3.4.2 shows somewhat more than a factor of five spread between the values of high and low (best-fit) data estimates.

3.4.4 Additional Considerations Affecting Data Volume. - Since the beginning of the space program, there has been an evolutionary development towards higher data volumes and higher raw information rates from the basic sensors as the experiments are designed to conduct more and more detailed measurements. This evolutionary trend will certainly continue during the next generation of space flight and may be expected to include such items as on-board general purpose computers for pretransmission

*"Apollo Applications Program Experiments Data Flow Plan," Pauley, R.J., Bellcomm, Inc.

**This 28-day mission life is consistent with Bellcomm assumptions used in their data generation estimates.

data processing, automation of data acquisition stations, and more extensive real time data relay to the control centers. One result of these new techniques will be to push the experimental data transmission rates nearer to the limits of the available communications channels discussed in the preceding paragraphs. An equally important result will be the more efficient use of these available communications channels. Estimates have previously been made that in many cases up to 99 percent of the data gathered merely furnishes a background level against which to measure the effect of the occasional disturbance which is the object of the experiment. * As a result of real-time experiment control and preprocessing of data prior to transmission, it will be possible to greatly increase the concentration of valid data actually transmitted. In effect, this constitutes a prereduction of data on board the space craft, and thereby reduces the amount of reduction required during post-processing at the ground sites. As a result, the increase in reduced and final analyzed data volumes should be significantly greater than the communications channel-limited increase in raw data volumes.

For the most part, these are the types of changes which can be forecast for the immediate five year period. By using computers and additional control and checkout equipment, the data acquisition stations should be able to reduce the setup time for each satellite pass and hence utilize more time for actual data recording. Other than this, the basic communication limitation discussed in Section 3.4.3 will remain essentially the same. Beyond the immediate five year period, however, it appears that orbiting data relay satellites may come into use. ** When they do, they will provide full time coverage at previously unattainable data rates, and thus effectively remove the upper limit for data acquisition established by the earlier communication channels.

3.4.5 The Data Base. - The figure presented in Section 3.4.2 represents the best current estimates of the raw data volumes. In order to complete the data base estimate, some indication of reduced and final analyzed data volumes is also required. Discussions with NSSDC indicate that no concrete relationship exists between the raw data and its subsequent reduced and final analyzed versions. In some cases, the data volume may be reduced by an order of magnitude or more, while in other cases it may be increased by that amount. The assumption here is a one-to-one correlation between the raw data and the combined subsequent versions. Consequently, the high and low data base estimates are as shown in Table XII. It should be noted that these figures do not include any record redundancy. In the actual operating system, such redundancy will be required for master records, but not necessarily for copies of records stored at other facilities.

---

*Telemetry Processing for NASA Scientific Satellites", Habib, E. J., GSFC.
**"Advanced Space Information Systems, Ludwiz", G. H., GSFC, April 67.

TABLE XII
DIGITAL DATA ESTIMATES FOR THE AAPA DATA BASE

COMPOSITION OF DATA BASE	ESTIMATE OF DATA BASE (Bits)	
	Low (Best-fit)	High (Communications Limit)
Reduced Data Final Analyzed Data	2.2×10^{11}	1.2×10^{12}
Raw Data (digital) Reduced Data Final Analyzed Data	4.4×10^{11}	2.4×10^{12}

3.5 Data Usage Considerations

Thus far, only the characteristics of the data base itself have been considered. It is further required to examine the various circumstances surrounding the utilization of this data by the overall user community. The reason for doing this is to disclose, as early in the planning stage as possible, what type of user trends may be anticipated, which will be high usage and which low usage data, what peak access periods may be expected, etc. It is possible to construct a hypothetical data-user model to help analyze these various usage characteristics. Such a model is shown in Figure 4, where the AAP Experiment Type Classifications of Table VIII are matched with the regrouped S.I.C. major classifications of Section 2.0. The connecting lines represent the most likely data flow paths which may be expected to emanate from the archives to the users. The 'Relative Demand' descriptors for these usage paths represent an interpretative evaluation of the first-level user responses obtained during this study. The figures in the upper left-hand corner of the S.I.C. blocks indicate the number of U.S. companies, as tabulated from Poor's Register of Corporations, Directors and Executives, engaged in that particular area of manufacturing. Areas of overlap (i.e., companies involved in more than one major area of manufacturing) are ignored since the intent is to obtain a gross overview of the distribution of major U.S. industries as related to the profile of AAP experiments. Using this model in conjunction with the data base considerations made previously, some basic observations can be made.

3.5.1 Initial Usage Characteristics. - Especially during the initial phases of AAP, the demand for data by users will have a direct relation to the most recent data (i.e., the only data) acquired. Consequently, it is particularly meaningful to review the data accumulation characteristics for AAP flights 1A through 4 as plotted in Figures 5 and 6. These figures show that the initial leader in data volume is the Space Engineering category of experiments, which then levels off to be overtaken by both Space Science and Life Science experiments at AAP-2. Since Space Science is considered to fall totally under the responsibility of NSSDC, it will not be considered here. Of the remaining digital data, however, it should be noted that Life Science data constitutes

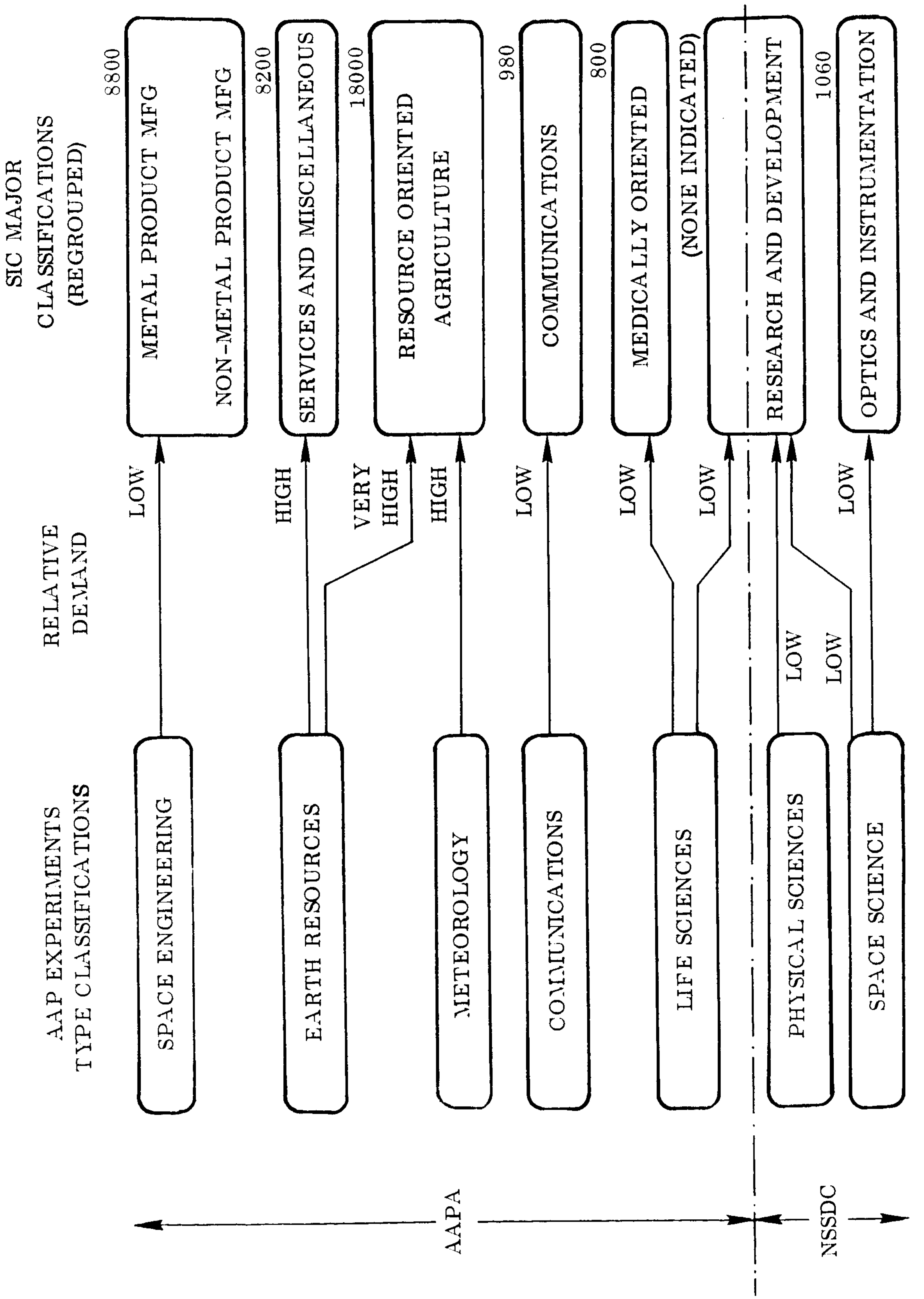


Figure 4 AAP Data Usage Profile

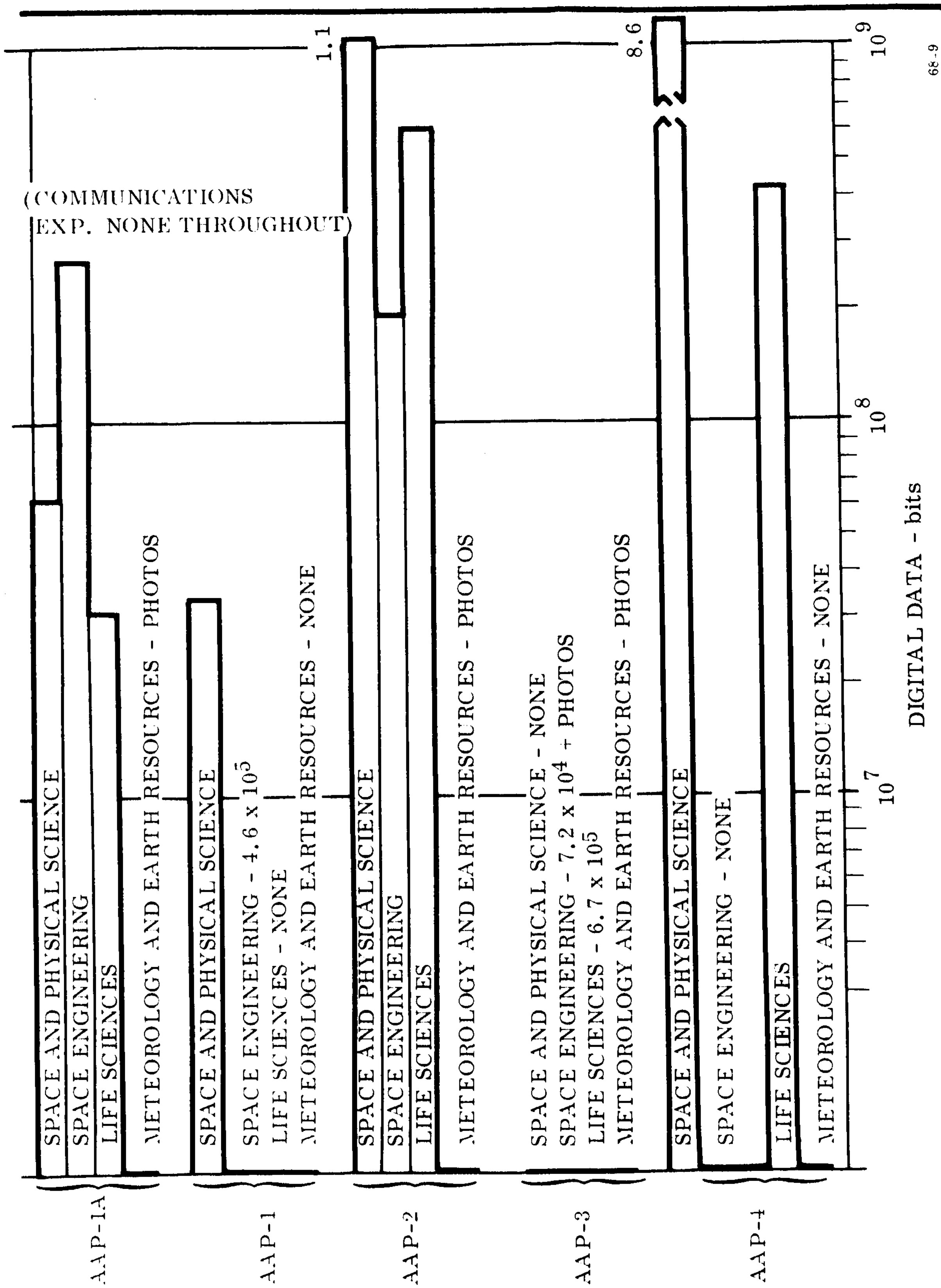


Figure 5 Data Volumes for AAP Flights 1A through 4

approximately 70 percent of the total digital data through AAP-4. Referring to Figure 4, it can be seen that both Space Engineering and Life Sciences have been classified as low usage categories. In fact, they may prove to be particularly low usage at first, pending the promulgation of initial PI results through the AAP publicity cycle. Only when the validity of the initial experiment results and some indication of their net worth to the user community are made know, will the AAPA begin to receive any significant volume of user queries. This vindicates the absolute necessity of having an effective publicity cycle even before having a full-fledged archive.

It can be further seen that, since the data so far discussed is of a digital nature, the archives can anticipate a significant 'lag time' between launch and recovery and the final receipt of data--raw, reduced or final analyzed-- at the archive facility. The duration of this lag time is dependent on the processing time required by both the NASA centers and the PI. It is a time impossible to stipulate and, at best, difficult to estimate. In most cases, it can be expected to be somewhere between one and two years (including PI investigation) although the real-time trends being fostered for AAP will certainly result in making much data available far sooner. During this initial time lag period, the AAPA should emphasize the development of its publicity cycle to the fullest.

One type of experiments data which is not shown in Figures 5 and 6 is photographic data. The reason for this is that not enough information is presently known about the photographic requirements of the various experiments to permit even a rudimentary estimation of image data volume. * Scheduled on flight 1A, however, are two earth resources experiments--S005 and S065-- whose principal mission is of a photographic nature. One characteristic of photographic data is that it does not require the intensive postprocessing required of telemetered digital data; it is available upon development. Another characteristic is that photographic data is its own best form of publicity. Also, it can be seen from Figure 4 that Earth Resources experiments can be expected to engender by far the highest usage rates. If AAP-1A is consummated as scheduled, there should be little lag time before the photographic data are made available to the AAPA, and a considerable influx of user requests may be expected almost simultaneously. This stipulates that the initial AAPA facility have available (or accessible) at least a moderately sized photographic reproduction capability.

3.5.2 Later Usage Considerations. - With the completion of a number of successful AAP launches, user requests can be expected to level off in a somewhat more predictable fashion. User data from previous experiments will have been gathered, and this can be applied to predict future requests for both past and future experiments of a similar nature. Since the highest usage data can still be expected to be Earth Resources and Meteorology, the logical trend will be the seeking of continual improvement

*Data Generated by Experiments on Flights 1A through 4", Agarwal, R.J., Sept. 22, 1967.

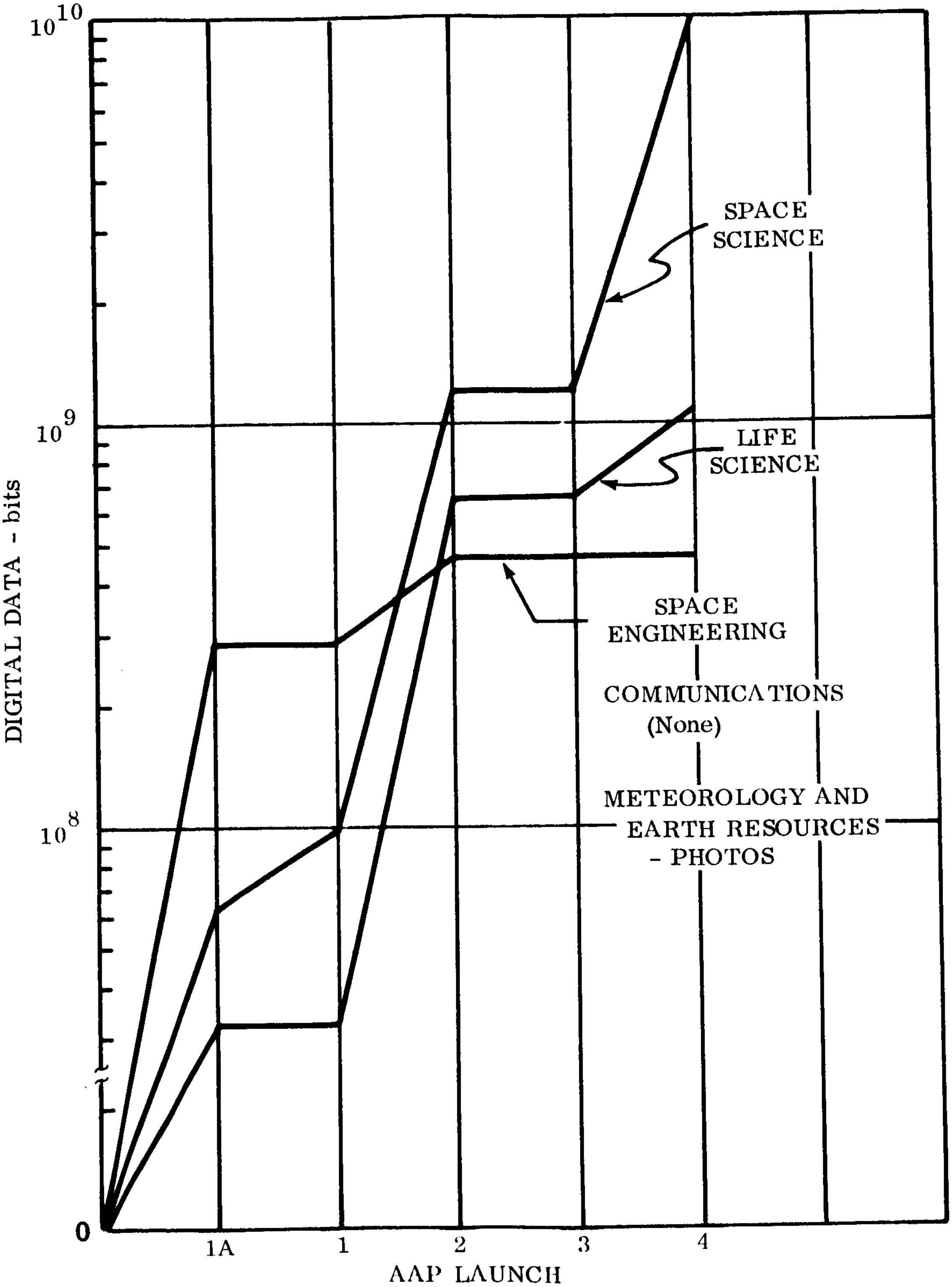


Figure 6 Data Accumulation Characteristics for AAP Flights 1A through 4

of user service in these areas. This can be primarily accomplished through the Dissemination Section which is discussed in Section 6.0. Dissemination is for the most part a production-oriented operation.

Even though the above tendency may be justified on the basis of user demand, there is question as to what degree it should be indulged at the possible expense of other available and useful data. The reference here is to the mass of experiments data contained in digital form. Although much of this may also contain a wealth of user-oriented information, it will require considerably more effort through the Surrogation and Announcement functions, discussed in Section 6.0, to make the availability of this information known. Since both of these functions are primarily of an intellectual nature, they tend to be more difficult, time-consuming and expensive to perform than the Dissemination function. Unless properly promoted by AAPA policy, much of this valuable, but less obvious, data may tend to lie fallow.

Even so, Figure 4 shows that it is entirely probable that much of the AAPA data will not experience high, if any, usage. As mentioned previously in this section, the most sensible archive structure for AAPA seems to be one which segregates active from less active data with eventual provision for data retirement to a more permanent archive, enabling the data base to be kept at a working level and more in line with storage and retrieval requirements.

3.6 Definitions

Examination of the overall nature of the data to be comprehended by any specific system configuration indicates that the data may be categorized into four major classes, each with distinct and major differences. The articulation of these classes appears to have useful implications for system design and to provide an appropriate opportunity for the establishement of a descriptive terminology. Since the four classes happen to parallel rather directly a similar terminology from scientific and scholarly research in general, that existing terminology has been adopted in the discussion which follows. The four classes, discussed in more detail below, because of the fundamental importance of the distinction involved, may be summarized here as follows:

- Primary data, typified by telemetric reduced telemetric, spacecraft stored, or photographic data

- Secondary data, typified by technical reports

- Tertiary data, typified by summary listings of experiments in progress, their cost, status, etc.

- Quaternary data, typified by summaries showing patterns in the tertiary data

3.6.1 Primary Data. - This the "raw or actual data brought back by sensor and recording systems. In many cases there appear to be several versions of it, each of which may have to be stored and managed. There are clean (reduced and partially analyzed) and dirty (raw telemetry) versions, analog and digital-converted versions, etc. The differences between these versions and between data records within the primary class are very large and problematical, but have relatively less immediate effect upon the process of conceptual design than do inter-class differences. Primary data tends to be voluminous, difficult and expensive to manipulate, and expensive or impossible to reproduce if all copies of its record were to be lost by system failure. It is clear that many of the possible AAPA configurations would involve the storage of much primary data in forms such as mag tapes, photographs, etc., where the record form must be dealt with not only as data but also as an object. There are, of course, other objects which the system must manage: objects such as the biological and soil samples and space debris mentioned in the original NASA work statement, and objects required in the system's own operations, for instance. The management of substantial collections of objects results in requirements entirely separate from the management of data, and these requirements may constitute one of the descriptive boundaries of the system.

3.6.2 Secondary Data. - This is usually derived from documents reporting the primary data. The most familiar form of secondary data is the formal summary, interpretation or commentary drawn up by the Principal Investigator (PI) and/or his staff and rendered as a 'published' technical report, conference paper or journal article. If the same primary data is reused by the same PI or, in a successful transfer of information, by another PI, more than one 'version' or families of versions of secondary data may occur. Strong provisions have been made elsewhere in NASA for the management of this data, under the direction of specific and considered federal and NASA policies. No justification has been found to consider the inclusion of the management of secondary data in these 'documentary' versions within the proposed AAPA. There are, however, a number of considerations involving secondary data which will have to be defined more exactly; some of these involve policy-type decisions where the considered judgment of NASA management may eventually have to to solicited. There is, for instance, the question of compatibility and commonality with other NASA and federal information systems; in many cases these appear to involve links between adjoining classes of data. How and to what extent will primary and tertiary data comprehended by the AAPA need to be linked to related tertiary and secondary data managed by, say, the NASA Scientific and Technical Information Center in College Park? Which agency will establish, store and manipulate the linkage information? The problem of definition here is not a minor one. There is, in addition, the question of alternative versions of secondary data not comprehended by the document-managing agencies. Some of these alternative versions may be particularly problematical because of the occurrence in a single record of two or more classes of data. Some examples include:

1) Lab notebooks and similar notes. Many laboratories have long had formal management arrangements for the storage and control of these, and the PI himself has a strong interest in them.

2) Backup material not included in the formal journal and technical-report versions. The existence of the material is frequently alluded to in the published versions, and is now normally found (when it can be found) in personal files or organizational files. The American Documentation Institute has sponsored a growing archive of such material, now housed and serviced by the Library of Congress. A number of professional and scientific societies are involved in similar archival activities.

3) Informal communications such as preprints, abstracts, notes for conference "talks", etc. Many professional and scientific societies are now heavily involved in the archiving and distribution to users of such communications. Their managerial arrangements vary from very good (AIAA) to disastrous. Involvement in such activities can be very expensive, even where the systems arrangements are efficient. NIH, for instance, was spending $416,000 per year on such activities in only seven rather narrow subject areas when its program was abruptly ended and NIH returned to its primary role as a research and research-sponsoring organization.

It appears unlikely at present that the AAPA system will store and manage significant amounts of secondary data; the relationship with other agencies which do handle it, however, will have to be formally specified, and will constitute one of the boundaries of the AAPA system. The specification of which types of linkages to secondary data will be handled by the AAPA system, and under what conditions, will describe one of the internal characteristics of the system.

3.6.3 Tertiary Data. - The most familiar form of this information, frequently called surrogate data, is probably the library catalog card; indexes and directories in general are composed largely of tertiary data. It is such information which is manipulated by the NASA/SCAN system; which appears in NASA's STAR publication and the indexes of the Apollo Parts Information Center (PRINCE/APIC) in Huntsville. Other examples might include the Executive Summary or "Manned Space Flight Experiment Catalog Listing" documents of the Manned Space Flight Program. Tertiary data may be generated by middlemen, or by middleman agencies, rather than by the Principal Investigator (PI) or the ultimate user. Attempts in the past several years to involve the PI in generation of tertiary data (viz. the use of DOD form 1473 in the back of DOD technical reports) have been energetic but largely unsuccessful.

3.6.4 Quarternary Data. - The most familiar form is the survey, usually based upon a manipulation of tertiary data, which groups and manipulates tertiary data to show higher order patterns. The manipulation may be performed to show, justify or establish funding patterns; review technical efforts by subject areas for program direction purposes. The AAPA system would probably include, as most such systems do, a provision for informing the searcher that his search query intersects an excessive number of tertiary records, and that he must add additional descripters to narrow the search; such an instruction would reflect quaternary data. The sophisticated type of bibliography librarians call a 'guide to the literature' usually involves quaternary data. The operation statistics which the system reports on its own operations are quaternary data.

3.7 First Level Flow Chart

The AAPA will be established by NASA to afford the widest practical distribution of non space oriented experimental data, both raw and reduced, resulting initially from the Apollo Applications Program. As such, it will provide an active repository of experimental data for all users. This section addresses itself to the functions of the archive and presents first level functional flow charts for a fully operational system. By the nature of the system and its data input, the complete AAPA discussed herein will gradually evolve from a smaller, more compact pilot system. The pilot system will be discussed in Section 3.11.

3.7.1 AAPA System Functions and Boundaries. - While this system has been called a data archive, its principle concern is with the dissemination of practical, current (nonspace-oriented) experimental data and not with the preservation of all AAP raw data. As such it will provide an active repository of operational experiments data for all investigators and will be responsible (Figure 7) for the active collection, organization, announcement, storage, retrieval and for the collection of feedback from users, and dissemination of data. The system must perform the functions of making known its usefulness, performing on users' requests, locating source data and conditioning this data for interpretation by a user. In the process of performing these functions, the AAPA will keep abreast of current user requirements through feedback from these users.

The process of conceptual design for a system such as the proposed Apollo Applications Program Archive which fulfills the aforementioned functions cannot proceed rationally and efficiently without a satisfactory and formal delineation of the boundaries of the system. Such a delineation is, in effect, the first and most general description of the system. It is also, initially, the description most useful in the early iterative phases of definitionof other system characteristics such as CPU requirements, memory capacity and structure, role of the ultimate user, I/O requirements, cost constraints, and the like. The boundary conditions are graphically shown in Figure 8. The General Archives System consists of the technological constraints of an input of data, an output of data, hardware, and software. These constraints must be determined and measured in terms of accomplishing the objective for which the system is to be designed and then tempered by the constraint of cost. The degree of decentralization becomes a secondary factor depending upon the requirements of the users and the trade-off of cost versus service rendered.

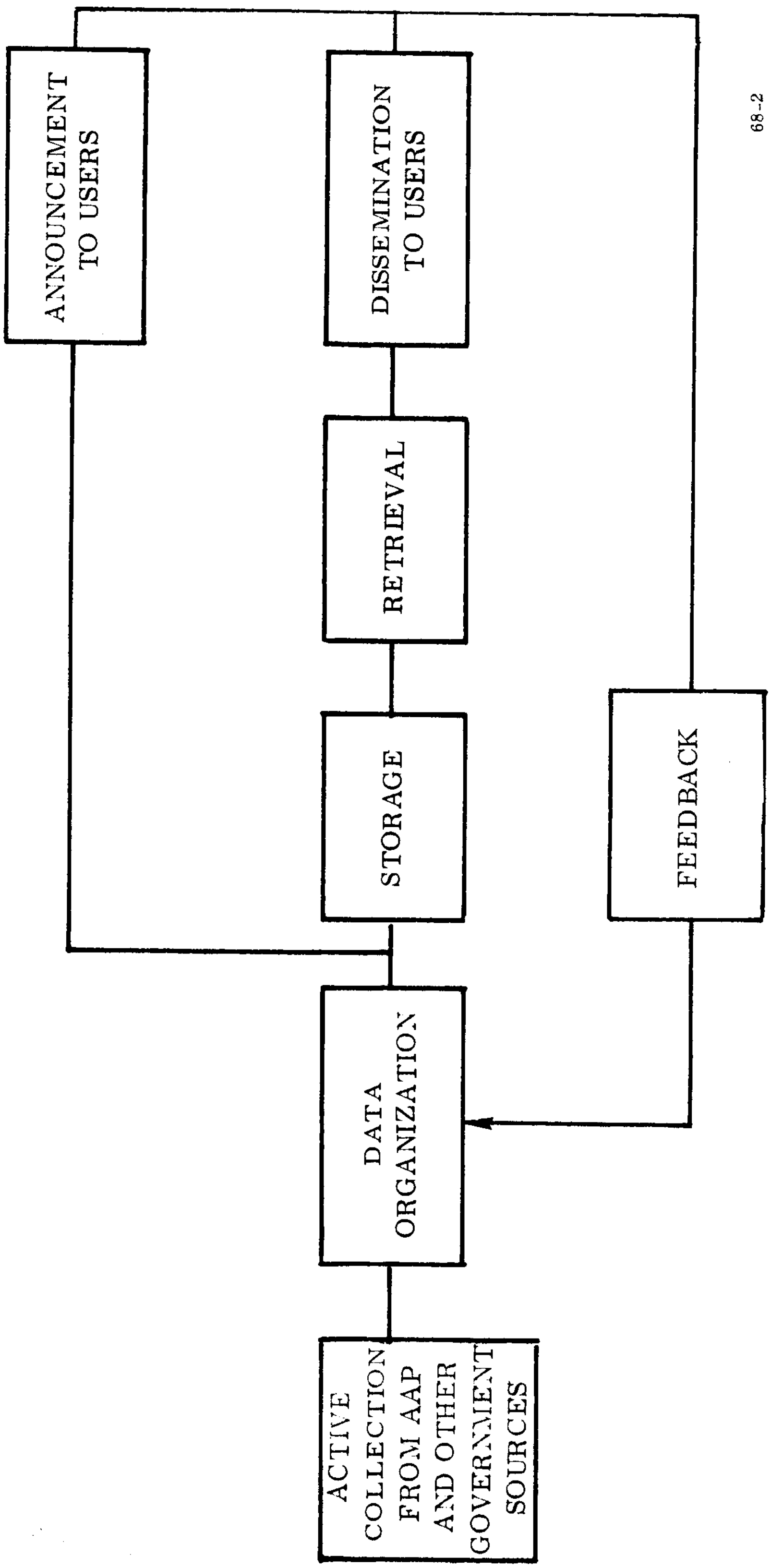
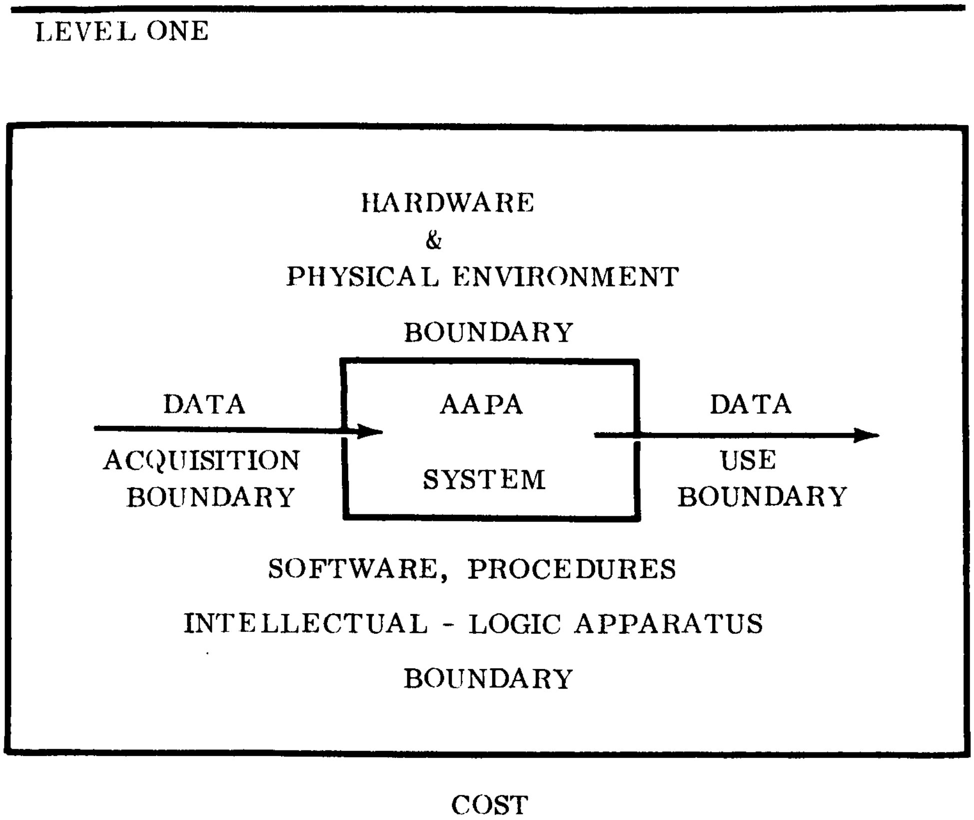


Figure 7 AAPA Responsibilities

GENERAL CHART OF FIVE
SYSTEM BOUNDARIES



67–2650

Figure 8 System Boundaries

3.7.2.9 Apollo Postflight Data Handling. - Since the hardware and support facilities developed for the Apollo Project are to be used to the maximum extent possible in the Apollo Applications Program, it is logical to first examine the mechanics of the postflight data flow of the Apollo Project. This has been done and reported by Bellcomm, Inc.* A suggested postflight data flow, shown in Figure 9, consists of the data handling activities from the time the data is first recorded on analog magnetic tapes at Manned Space Flight Network (MSFN) tracking sites until the time these tapes are placed in permanent archival storage.

Tapes from all of the MSFN sites except Texas (TEX), Guaymas (GYM), and Cape Kennedy (MIL) are transported by commercial air transport to the NASA/GSFC, Manned Flight Operations Division. At Goddard, copies of the tapes containing launch vehicle data are made and forwarded to Huntsville, Missions Operation Office; copies of a limited number of tapes are made and retained by Goddard network engineering personnel; and copies of selected tapes are made and forwarded to various contractors as specified by the Program Support Requirement Document. The bulk of the original tapes are forwarded to Houston, the Computation and Analysis Division, for analyses; however, some selected original tapes are also forwarded to the Space Division of North American Aviation for analysis.

Tapes from the remaining three MSFN sites - Texas, Guaymas and Cape Kennedy- are transported by commercial aircraft directly to Houston's Computation and Analysis Division. Here, copies of the tapes containing launch vehicle data are sent to the Mission Operations Office in Huntsville; copies of a limited number of tapes are made upon request and forwarded to the Manned Flight Operations Division of Goddard; and copies of selected tapes are forwarded to the various contractors as specified in the Program Support Requirement Document. Again, some original tapes are forwarded to the Space Division of North American Aviation for analysis and the remainder (the bulk of the original tapes) are retained at Houston.

When North American has completed their analysis, they return the original tapes to Houston. After two missions or twelve months, all of the original tapes are then sent to Goddard for permanent archival storage.

3.7.3 AAPA Input Data Flow. - Based upon the Bellcomm report of the Apollo postflight data flow, a logical input to the AAPA for flight data would be similar to that shown in Figure 9. It should be noted that AAP data extracted from MFSC telemetry tapes go directly to MSFC from GSFC. Likewise, AAP data extracted from MSC telemetry tapes goes directly to MSC from GSFC. Then, since all original tapes will be routed to their respective control centers, it is assumed that these centers will be responsible for the decommutation process of telemetry data and for the distribution of the decommutated copy tapes and hard copy data (photographs, etc.) to the experimenter and the

*"Apollo Data Flow", Case 600-2, Pauley, R. J., Bellcomm, Inc., 13 July 1967.

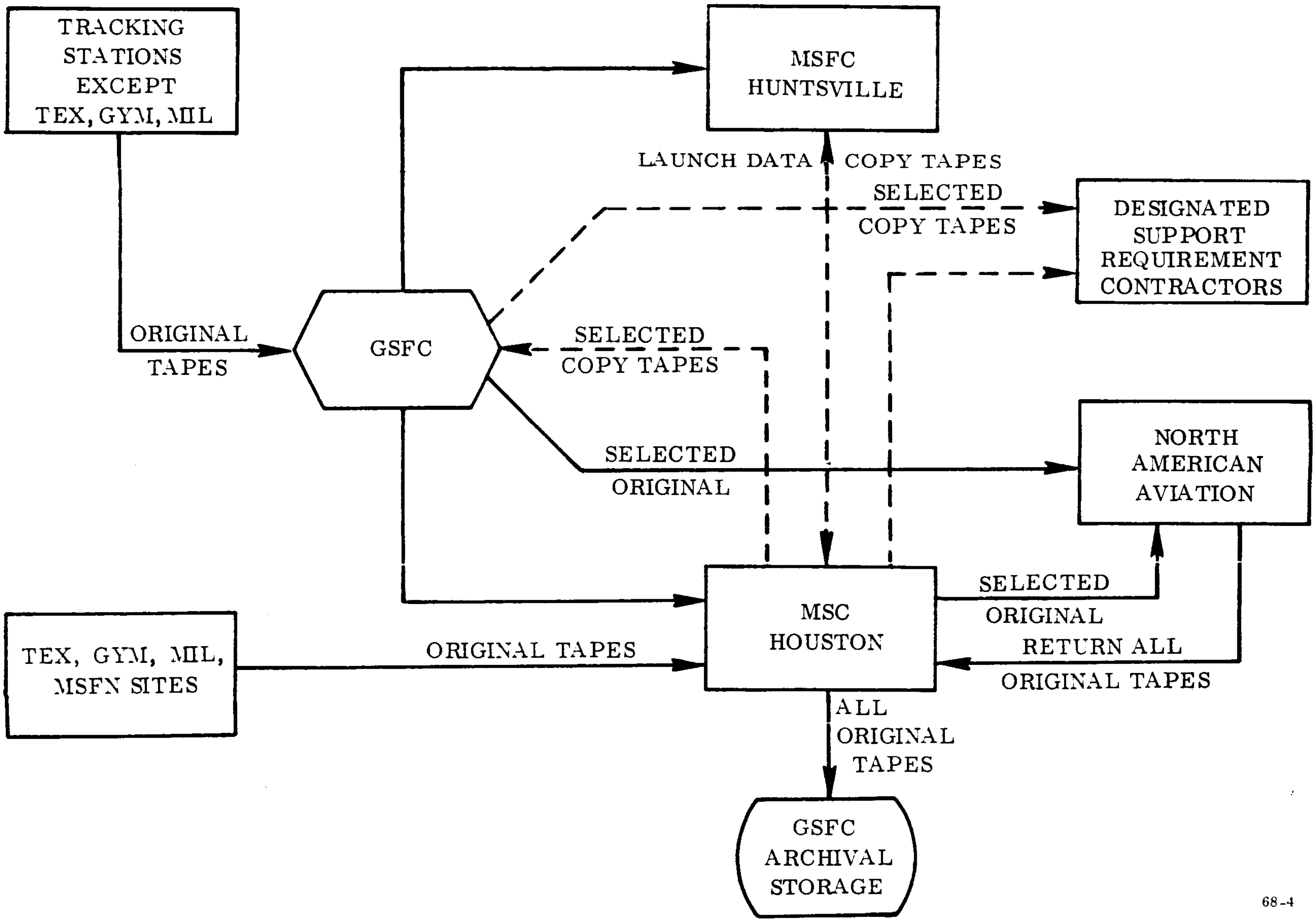


Figure 9 Postflight Data Flow (Present Configuration as Reported by BellComm)

AAPA. This in no way shifts the responsibility or the point of contact from the other NASA field offices as far as the experimenter is concerned. An experimenter must still deal through "his" field office, and said field office would be advised by the control center when data is forwarded to one of its experimenters.

A diagram of the complete AAPA input data flow is given in Figure 10, including both data flow and control paths. In discussing the system input, the data flow will be considered first and then the necessary "controls". A more complete discussion of the overall control considerations will be rendered in the following sections. Figure 10 has been composed based upon the assumption stated above. The alternate assumption could have been made that each cognizant data center will perform the decommutation of data for its own experiments - receiving the raw analog tapes from Houston. The use of one or the other assumption is of little consequence to the AAPA system and is included merely for completeness.

Several types of tapes are candidates for storage in the AAPA, including 1) reduced analysis (reduced data and final analyzed data) tapes returned from the experimenters; 2) extract tapes for the experimenters; 3) selected data obtained from secondary input sources; and 4) various digital tapes that can be produced from the analog flight tapes. The first three types would compose the active data base referred to earlier while the fourth would generally be in an inactive data base. The secondary input sources referred to in the type 2 data include many of the data centers currently in existence which store various types of Earth related data. Thus, should the AAPA receive a sufficiently large number of requests for data which are stored at one of these other data centers, it may be desirable to obtain and store copies of such data at the AAPA. A Data Acquisition Group will be responsible for the actual physical acquisition of all data.

Prompt and complete acquisition of data must be guaranteed for the system if it is to operate effectively. Such acquisition cannot be assumed and must be secured by positive and deliberate efforts. Acquisitions Officers must be provided to expedite and ensure data acquisition and some of these must be remotely located (regardless of the location of the AAPA control staff and main archive) with adequate travel funds. The Acquisitions Officers must acquire tertiary data for the system as well as primary data and over a long chronological span from before the AAPA flight to after the completion of the experimenter's contract. Acquisition Officers will have to have considerable familiarity not only with experimental data techniques but also with the NASA project management process.

Once data has been received by the Data Acquisition Group, it will be forwarded to a data control department which will consist of a data control group and a use analysis and technical writing group. The Data Control Group will be responsible for the generation of most of the tertiary data associated with the data base, for any data file conversion, and for the storage of (digital) catalog and abstract data. Tertiary data commonly includes such things as library catalogs (cards or bound documents), indexes,

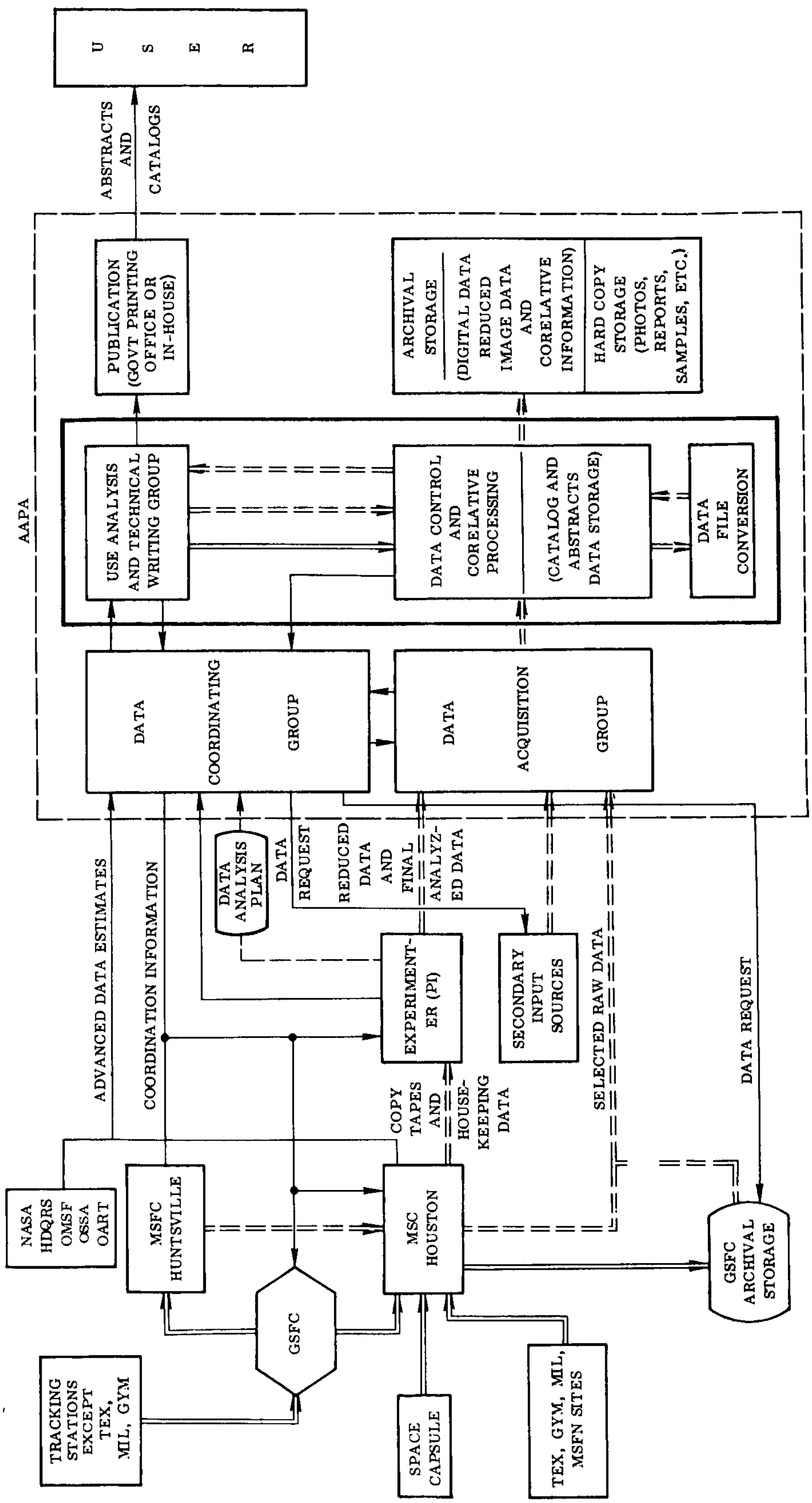


Figure 10 AAPA Input Data Flow

cross referencing and directories; however, the tertiary data generated for the inactive data base will be limited to the minimum needed for indexing.

The AAPA Catalog should be partitioned into small systems oriented to the needs and terminology of particular groups of users. This must be paralleled by a partitioning by specialty of the data use staff. Each of the specialists must be responsive enough to the needs of his own list of customers to effectively prevent him from generalizing. The actual mix of backgrounds of these specialists may be determined by the system administrator. It is expected that at least four specialists will be required at the outset of system operation.

The specialists should be centrally based with adequate travel funds. In some cases, assembling users at AAPA expense may be preferable to having the AAPA data use specialists travel, and travel funds should be available for such purposes.

The AAPA Catalog products must be responsive both in content and organization to the needs of the data use specialists in their dealings with potential users.

As mentioned above, one of the functions of the Data Control Group will be the performance of any data file conversion. The purpose of the data file conversion process is to convert the various data to be archived into a media and format which is more amenable to the AAPA Storage and Retrieval System. This will include the reformating of digital data into a common system format, with the addition of any necessary indexing tasks for data identification and retrieval purposes. It will also entail the transference of image data (mostly pictorial and graphical) onto either a film or magnetic base which is physically less bulky and lends itself more readily to system manipulation. In the latter case, where the data transfer involves image size reduction, there results a loss of resolution, and hence a loss of potentially recoverable information. Therefore, it will also be necessary to retain the original hard copy version of the image data in permanent though a less readily accessible archival storage. Implicit here is the incorporation of sufficient cross-indexing information to allow the original hard copy item to be located from the "compressed image" version, and vice versa. (Since the compressed version of the image data has less resolution and hence contains less data, it may be instructive to think of the compressed version as an "abstract" of the original. This is especially so in light of the concept of primary, secondary, tertiary, and quaternary data mentioned elsewhere.) When the input processing has been completed, the data will then be placed in "archival" storage for ready access to user requests.

The control or coordination of the input data flow rests primarily with the Data Coordinating Group. This group will interface with the NASA field offices, the experimenters and other data centers externally and with the data control department internally. The NASA field offices 1) will furnish preliminary data estimates to assist in advanced system planning; 2) will serve as the initial interface between the AAPA and the experimenter; and 3) will furnish the archives with copies of selected raw data tapes that users may request.

Similar to the procedure employed by the National Space Science Data Center (NSSDC), it is believed that the experimenter will be under a contractual obligation to NASA to furnish his reduced data to the AAPA. Some fixed time following the receipt of his raw data from Houston, the experimenter will be required to file a data analysis plan to the Data Coordinating Group which specifies in what stages he will reduce his data and what he anticipates furnishing to the AAPA. From this plan, definite agreement can be reached as to what data the experimenter will furnish and what time schedule he must follow in submitting the data. The AAPA will, in general, be most interested in receiving reduced data, namely the raw data from Houston which has been calibrated, correlated, processed to remove noise, and is in a useful form (as a function of time, distance etc.) to undergo analysis; and final analyzed data which is that data which the experimenter uses to support his findings.

Internally, the Data Coordinating Group will be responsible for overseeing the data flow operation. To this end, group personnel will follow the data through its processing, file conversion and storage, and the group will maintain copies of all catalogs and indices. As mentioned above, the primary interface in this task will be with the data control department, which consists of a Data Control Group and a Use Analysis and Technical Writing Group.

The Use Analysis and Technical Writing Group will work closely with the Data Control Group on the indexing task. The group will consist of competent technical people who are capable of analyzing the best potential uses of data to the extent of being able to determine the optimal "tags" (index designations) to assign to a given block of data. The indexing task could prove a fairly complex problem because of the various applications possible for a given block of data. Photographs, for example, could be categorized by the latitude and longitude of the photographed area; by the geological, oceanographic, or meteorological phenomena photographed; by the time of day at which the photograph was taken; by the camera used intaking the photograph, etc. Thus, members of the group must be able to discern all of the possible ramifications of a given block of data.

For example, Professor A. J. McNair of Cornell University who has had considerable experience in the area of photo applications, has provided a list of the surrogate information which his staff deems essential for referencing photographs. These reference items are as follows:

1) General location of exposure station - this can be given in longitude and latitude or in universal trans-Mercator projection.

2) Altitude of exposure station - this parameter, in conjunction with item No. 1, would provide X Y Z, location of the exposure station at the instance of exposure for each photograph.

3) In some instances, items 1 and 2 could be satisfied by a complete definition of orbital parameters. Orbital data might be preferred if adequate computer facilities are available.

4) Orientation elements - these would consist of the direction cosines of the camera axis with respect to the geocentric center of the earth. In other words, the axis of the camera would be defined in terms of geocentric earth coordinates.

5) Camera parameters -

   a) Focal length of the lens
   b) Speed of shutter
   c) Type of film used
   d) Kind of filter used
   e) Type of lens
   f) Calibration characteristics of lens
   g) Lens distortion characteristics
   h) Fluidicial marks

6) The exact time and date on which the photo was taken - time should be supplied in the most usable form and this would probably be Greenwich Mean Time.

The above items actually constitute requirements for the use of space photographs in the fields of photogrammetry, photo interpretation, and geodesy. However, if this parametric data is to be supplied, then action must be taken while the data handling plan for the AAP is still in its formulative stages.

The members of the Use Analysis and Technical Writing Group must also include competent technical writers who can prepare abstracts for the AAPA catalog (to be used for data searches) and prepare something equivalent to the "Data Users' Notes" employed by NSSDC. Included in these notes are the following: a description of the experiment and the measurements taken; the background of the experiment and the experimenter; the data reduction technique employed; the format or formats of the available data; a bibliography; and a list of references. The cooperation of the experimenter would be expected for validating such notes. The Use Analysis and Technical Writing Group will also be responsible for the actual preparation of a catalog for general distribution much like the STAR catalog or NSSDC's data catalogs "Satellite and Rocket Experiments" and "Catalog of Correlative Date". This textual material may then be printed either in-house or by the Government Printing Office.

3.7.4 User Access and Retrieval. - After reviewing the various promulgatory, or tertiary, data concerning the nature of the experiments data stored in the AAPA, the potential data user is in a position to make a preliminary evaluation as to which data he desires, and initiate an inquiry or request to the AAPA. The terms inquiry and request as used here, are differentiated as follows: an inquiry connotes a query for further information concerning the archived data, while a request connotes an actual request for data retrieval, either for on-site review or distribution to the user.

It is assumed that in spite of the descriptive tertiary information provided as part of the AAPA publicity cycle, the user will probably not have sufficient information at hand to decide conclusively which data he requires. Consequently, the user should have the benefit of successive interrogation loops into the archive before actual retrieval and distribution is effected. For example, assume the initial user inquiry is in somewhat generalized vein. The AAPA would then respond with more complete tertiary information, on the order of more abstract information, catalog listings, etc. After reviewing these, the user might then interrogate the archives with a more specific inquiry, resulting in more specific tertiary information being output for his review. Once he feels he has enough information - including the type of data, media, quantities, and prices involved to make a decision, he may reenter the retrieval loop with a data request. This may be a final request, resulting directly in a release of data for distribution to him, or at his discretion and convenience may include a final onsite review of the data he has requested.

Figure 11 shows both inquiry and request loops as they would exist in the AAPA. Both would be initially processed through the Data Coordination Group, since this group retains overall responsibility of all archival activities. The actual processing of the inquiries and requests would probably be combined as well; however, these functions are separated in the figure for purposes of clarification. A fundamental difference between the two processing functions might be the actual data retrieval: the tertiary data, which would be accessed by a data inquiry, would probably be stored in a smaller sized, more accessible storage media; whereas the bulk archival data, which would be accessed by a data request, would be stored in the archives per se.

Following machine retrieval of inquiry type daya, compilation is required to prepare the assembled data for distribution. This consists primarily of converting the electronically stored data into hardcopy - primarily unbound paper reports. The nature and relatively small volume of this operation should enable it to be accomplished by AAPA personnel.

The retrieval and compilation of request data will involve the handling of much larger quantities of data and types of storage media. After machine location of the stored material, archive's personnel will have to retrieve and handle such diverse and incompatible media as magnetic tape, microfilm, hardcopy photographs, and document hardcopy. As previously mentioned, once compiled, this data may merely be prepared for distribution if desired by the requestor. However, due to bulk of the data and appreciable cost of the media involved, the user should also have the option of an onsite review if desired. This gives one final check of the data before release for distribution, and also provides another dimension for utilization - that or potentially costfree use of

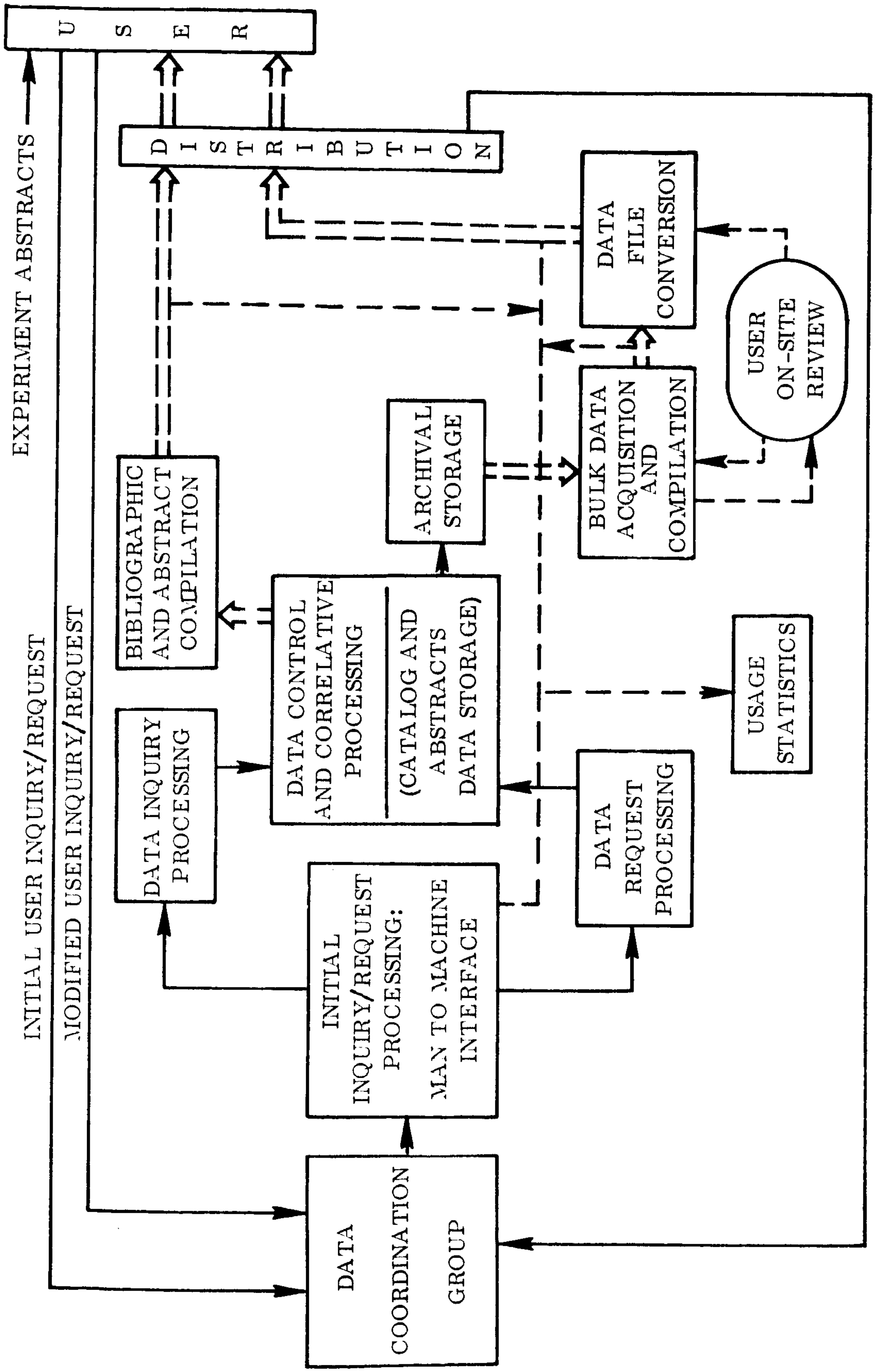


Figure 11 User Access and Retrieval Data Flow

the archives for users who would be perhaps unable to justify funds for data procurement* but are able to spend time onsite for perusal of the archived data.

One final function - that of data file conversion - will be required if the archives are to be beneficial to a reasonably broad scope of users. The media, codes and languages of data storage within the archives are not necessarily those being used by the users' data processing facilities. Also, the user with a specialized, or more modest installation will probably not have the required means of data file conversion at hand. Of course, for the archives facility to attempt to "custom fit" outgoing data for each user would be prohibitively costly; nevertheless, it is within the realm of practicality for the archives to perform a limited degree of file conversion so that a greater number of potential users can be more effectively served.

Data distribution, both for an inquiry and request data, can probably be performed as a mail-room operation for the initial configuration. Following the growth of both the archives and its usage, this function will probably expand to the extent where it could be assembled, by a separate distribution agency.

3.7.5 System Update Concept. - As indicated in Figure 11 usage statistics, a form of quarternary data, will be gathered about the data inquiries, the available data found by a data search, the data requested, and the data selected after a review of all data by the user. These statistics will then be combined for use in the system update concept shown in Figure 12. The crux of the system update concept is the operational review, a computerized analysis of the statistical data. From this review, indications can be output concerning advanced system planning; modifications in the storage and retrieval concept such as thesaurus update, subject index modification and cross referencing modifications; and modifications in the data base such as expansion, retirement, format modification, etc. The operational review will be under the supervision of the Data Coordinating Group who will also have control over modifications in the data base - in particular data retirement.

It is anticipated that when data is retired from the AAPA, it will be placed in permanent archival storage in a facility such as a local Federal Records Center. At such time, it will also be necessary to update the AAPA data catalog and compose a catalog of retired data to be maintained in the AAPA archival storage. The responsibility for the modifications to the catalog and for the maintainance of the retired data catalog would fall to the Data Control Group.

3.7.6 Coordination of Data Centers. - As mentioned in Section 4.0 above, there are several data centers presently in existence which deal with or are devoted to some aspect of Earth resources data. Among these are the National Weather Record Center,

*Usage costs, as such, would be required primarily to cover the cost of expendables required for user distribution, i.e., magnetic tape, photographs, etc.

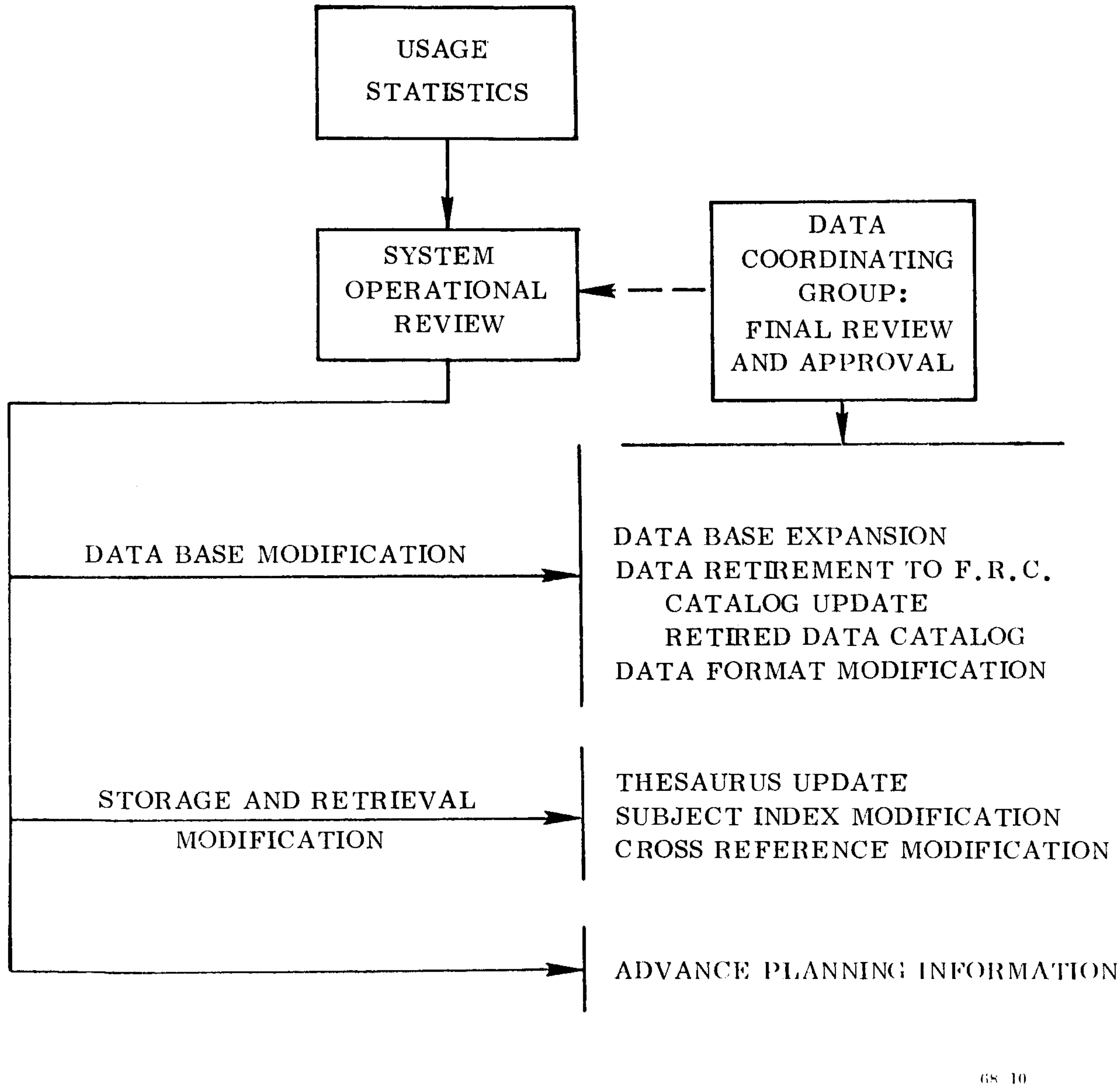


Figure 12 AAPA System Update Concept

the National Oceanographic Data Center, NASA's Scientific and Technical Information Facility, NASA's Earth Resources Data Center, the Coast and Geodedic Survey Library, and the National Aeronomy and Space Data Center. It is therefore desirable to establish and maintain some form of coordination between these data centers. One means of accomplishing this, as represented in Figure 13,is to have one person on the staff of each facility cognizant in a somewhat more than general sense of what data and/or information is stored in each of the other facilities. In the case of the AAPA this would be someone in the Data Coordinating Section. Thus, when a user comes to the AAPA, for example, with a request for data that the AAPA does not have, the AAPA could direct the user to that facility which would most probably be able to fill the request.

Figure 13 contains a block for NASA's Office of Technology Utilization (OTU), because it is anticipated that in many instances a user may not know exactly what data he requires or what data is available to solve his problem. In such a case, he could deal through a local Technology Application Center (TAC) office rather than traveling across the country to review data at the AAPA facility. For small businessmen, where the cost of such travel could be prohibitive, the use of Technology Application Centers as a local input/output would serve as an additional asset. Thus, users of the AAPA have two possible interfaces: 1) directly with the archives' Data Coordinating Group or 2) on a local level with the Office of Technology Utilization. It is to be emphasized here that the OTU is not being restricted to operating solely as a local I/O and data search mechanism for the AAPA. These are added tasks which can be easily assimilated with their present functions.

3.7.7 <u>Preflight Announcement Cycle.</u> - To best serve its function as a data "library", it is desirable for the AAPA to have some advanced feeling for what data will be most useful and most desired by potential users. One means of accomplishing this is through a preflight briefing and survey cycle as represented in Figure 14. In this cycle, the NASA field office would notify the AAPA of all experiments and anticipated data as the experiments are approved. The Use Analysis and Technical Writing Group of the AAPA could then, with the cooperation of the experimenter if necessary, prepare abstracts of these experiments and anticipated data, and distribute them to potential users. The users would inturn be able to assess their anticipated requirements for data and furnish estimates to the AAPA.

As a byproduct of this process, users would also be able to make recommendations through the AAPA for follow-on experiments. An example of such modification might be an experiment wherein a sensor is to gather data from a particular area of the Earth. To a given user it may be desirable to have data also provided from another area, and he could request this modification. In this capacity, the AAPA would act as a filter between NASA and the user, thereby permitting feedback to NASA on the experiments, yet limiting the influence and knowledge of the user on internal NASA functioning or direct intervention and interference thereby causing unnecessary work.

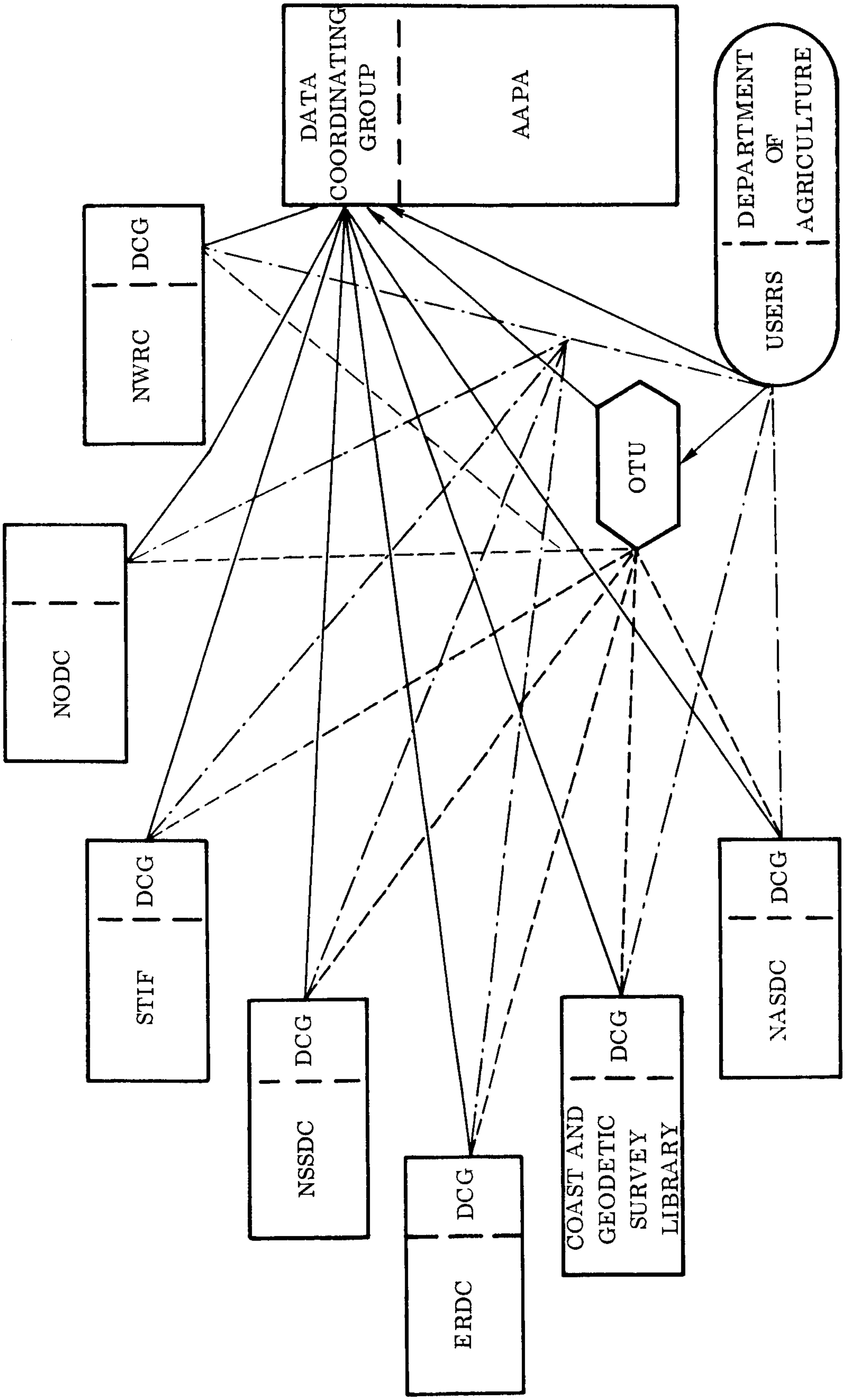


Figure 13 Coordination of Data Center

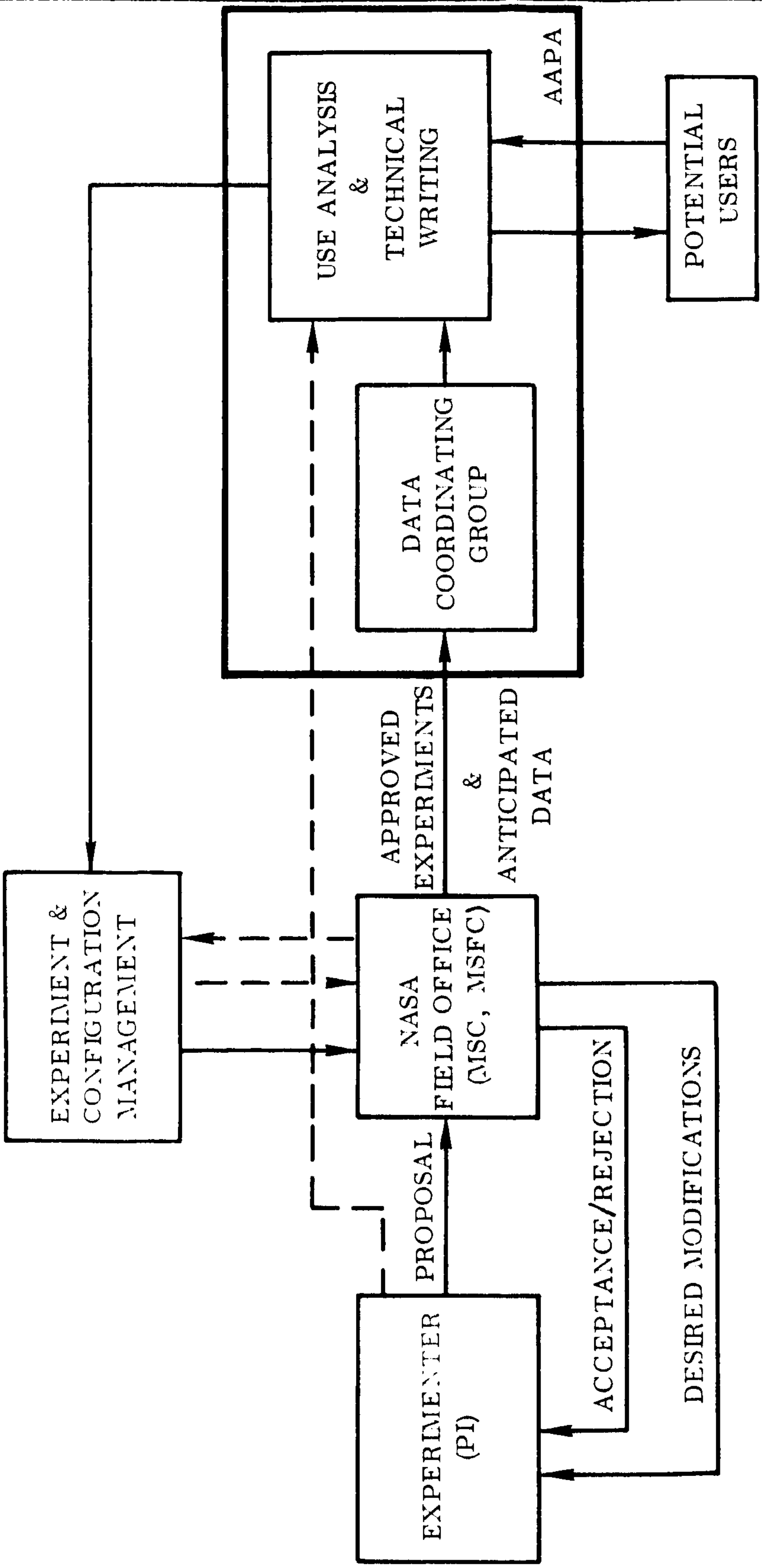


Figure 14 Idealized Preflight Experiment Publicity Cycle

3.8 Control Considerations

Assessment of the realities of the data flow in and out of the AAPA shows a pattern with strong implications for system design. Data flows to the archive, if it does, through a complex process involving a large number of widely scattered people and organizations whose attention is chiefly devoted to other activities. The data itself is of extremely diverse content and in formats so various and numerous that the types of record media to be handled are not even known in advance. Data use is by a geographically dispersed group of users who in total cannot be identified in advance with certainty and who have no organization, subject or discipline in common.

It is a messy situation and one which is potentially quite expensive, or disastrous, or both for the AAPA. Effective steps to manage such a difficult situation will require immediate action to create a broad and facile information tool to support the AAPA data control subsystem.

Specifications follow. It should be noted that they are specifications for the whole data control subsystem not merely the automated record keeping part of it.

3.8.1 Specifications. - The essential features of the data control subsystem must be capable of being implemented no later than the first quarter of the first year of AAPA operation. The span of its control must cover data in existence and in the archive, data packages not yet in the archive, and even data planned but not yet in existence. The span of control may also have to include data projects in the proposal stage but not yet scheduled; however, handling such data does not appear to be a requirement at this stage. Control must be exerted over raw data within AAPA's charter even if not archived in the AAPA stores. Data in the data control subsystem must be capable of being made widely available in a variety of forms and packages appropriate to the use and must be available via both immediate access and belated access techniques.

The data control subsystem must interface easily and effectively with acquisitions operations, with storage/retrieval and use activities, and with the data use promotion program as well as with the data control function itself. The subsystem must be capable of accepting inputs quickly and easily over each of these interfaces and in effect serving as the principal record keeping tool of each.

Prospective users of the data control system must find the effective subject structure of its products and access tools shaped to meet the needs, terminology and actual patterns of use of the customer. The single data control system should appear to the user as if it were in effect a small specialized system substantially tailored to the user or the user group of which he is a part.

The data control subsystem must include 'adminstrative' data elements (dealing with the identity of the data; e.g., principal investigator, contract number, flight designation, contract sponsor, etc.) as well as content data elements. Administrative data elements must be included broadly enough and must be manipulated by the system so as to make unnecessary any other major AAPA paperwork system.

It is reasonable to assume that the experimental data output will be greater than the input if this is to be an operating archive. Therefore the "housekeeping" data base will increase the total number of data elements which must be estimated. In other data control systems for this type of application an increase in effective data base can be as much as 50 percent. A reasonable estimate for AAPA is 30 percent increase in data base if consideration is also given to generating feedback information to NASA as to the effectiveness of this program. The programming requirements is a factor of no small concern. The executive and subroutines have not been identified as they are beyond the scope of this contract but estimates of thirty ($30.00) per instruction are common.

3.8.2 Implementation. - This specification recognized the political and technical need to get the AAPA system producing immediately. The specification is thereby nearly axiomatic. Early implementation does not require immediate operation of all of the tools and services of the AAPA, but a core of the data control routines and many of the data use specialist's tools must be operational by the end of the third month or expensive substitute systems will appear. The need for early implementation, like several other factors, suggests adaption of the existing compilation facilities and programs of similar agencies and use of existing information networks rather than an attempt to synthesize an entirely new operating system. There are several existing systems similar enough to the proposed AAPA system to allow successful adaptions. A proposed budget for the initial system, presented in Section 3.11.3 allocates money for both internal and external systems work as well as machine time purchases in response to this specification.

3.8.2.1 Chronological Scope. - This specification recognizes the common need to manage data not yet in hand, a need which has resulted elsewhere in the large data management systems of the Air Force, in the experiment-management systems of NASA and DOD and other agencies, and in even more catholic activities such as the Smithsonian's Scientific Information Exchange and NBS Standard Reference Data System. Some of these schemes are user oriented, some are intended to serve the management process itself; AAPA's provisions for broad coverage of differing project statuses would serve both ends. Since a strong and effective acquisitions program is essential to the fruitful functioning of the AAPA, the implementation of reporting procedures for projects pending must occur very early in the implementation process.

Arrangements for the support of the acquisitions function may be satisfied in a relatively simple manner through the design and use of a reporting form for use by the Acquisition Officers to be transmitted either by air mail or teletype. These forms would usually be filled out by the Acquisition Officer himself rather than by the investigator or a project related person and would cover not only the initial record creation task but the recording of status updates and data additions as well, thereby substantially assisting the work of the data control function. The Acquisitions Officers would be assisted by the provision of standard input data manuals and the AAPA index.

3.8.2.2 Availability. - Definitions of the data base of the data control subsystem must be widely available not only to the data control group at AAPA quarters but also in the field to the Acquisitions Officers, to data use specialists, and to users themselves. The number and variability of points of use suggests that access should be through printed and computer printout products rather than direct access, especially during the first few years of operation; a conclusion which is also suggested by the relatively small size of the defined data control data base. The AAPA Index which is proposed to meet this need would be available in printed editions for wide distribution and would appear regularly (e.g., monthly). A computer-printer edition for use by staff only would appear more frequently (e.g., weekly). Each would have the same data arranged in a number of different files (indexes) by appropriate search keys; e.g., geo graphical area, expected data of availability, location of supplemental data in the AAPA Catalog, etc. A non-index section in each edition would summarize separately the changes and additions since the last edition.

The data used to prepare the AAPA Index would be input by economical tape processing methods and could be searched economically by data use specialists using the same techniques. As an alternative procedure, new surrogate data for the data control subsystem could be input through paper tape punching typewriters also used to produce a manual card file. This technique would not satisfactorily meet the data needs of the Data Acquisition Officers in the field and is not recommended.

3.8.2.3 Interfaces. - If strong and effective arrangements are made to reach potential data users, the data may still be slow and disorderly in its processing because of ineffective interfacing of use procedures with control and acquisition procedures. A particularly critical interface is the acquisition interface, and it is apparent that the Acquisition Officers must become involved in the acquisition at secondary and tertiary data as well as the primary data itself; indeed recently designed project-management schemes such as the Navy's NARDIS system now normally include such a technique. In document work, the DOD form 1473 system is essentially a similar technique created to meet similar needs.

The AAPA system will frequently have to interface with associated systems for input and output purposes, and the intellectual design of its classification and subject indexing, data elements schemes, etc. must be developed compatibly with these systems. Particularly important will be the NASA project management systems and the NASA/DOCINC document control system.

3.8.2.4 Subject Structure. - The AAPA will have many of the characteristics of a large system even in the pilot stage. It is characteristic of large systems that they tend rapidly to become unresponsive to the needs of user subgroups and hence to breakdown. There are strong theoretical reasons underlying the phenomenon (see A. D. Little's

1963 report, Centralization and Documentation) and a mass of practical evidence of its existence. Provisions to counter the phenomenon are available but tend to run somewhat counter to the need for a strong central control tool. Two of the essential tools are discussed below.

The first of these tools is a mission oriented staff of data use specialists oriented by background and assignment to work directly with particular industries and industry groups, in sufficient number and experience, and with liberal travel funds. A regional distribution of generalized field officers does not meet the specification and will be excessively expensive and ineffective. Regional representation may be effected in part by a nonparticipant dissemination of printed indexes and advertising material through the existing network of technical information centers, information analysis centers, company special libraries and, of course, other data archives.

The second tool is the AAPA catalog schemes. This will be tied into the AAPA Index data base but produced and organized independently. Its modules will give much deeper information about the content and format of the data packages both in and anticipated by the archive, and must be capable of being disseminated and maintained independently of the whole catalog. The abstracting and surrogation techniques used will be slanted to the potential user. It must be anticipated that considerable innovation will have to be employed in evolving appropriate surrogation techniques for the AAPA Catalog. Neither the NSSDC data briefs nor the familiar abstracting and indexing techniques of document work will be useful models.

3.8.2.5 Data Element Coverage. - Examination of comparable information systems indicates a span covering more than 100 data elements is extremely common with a range from the 400 of the OAR project management system to the approximate 100 plus of cartographic management systems. The AAPA system should be able to handle at least two hundred, with maximum lengths of 5000 characters for tertiary records, and a substantial number of variable length fields. Because the character set must be extensive to be responsive to the normal reading habits of users and because output printing must be of high quality, a nine-bit 256-character set of characters must be accommodated for internal coding purposes; although narrower sets may be used for many input and output purposes and narrower coding patterns may be used for some system data stores.

3.9 Software Consideration

The AAPA data base will be fairly unique in content and cannot be thought of in the terms commonly used for information retrieval systems. Information retrieval systems are normally a collection of facts which are arranged in a computer system storage such that they can be retrieved or compared to a given set of criteria or search parameters. The AAPA data will not represent individual useful facts but rather data streams collected from various sensors aboard space craft. It is anticipated that the smallest retrieval unit will be all of the data stream from a particular sensor and experiment. A further subdivision may be possible which is a time period subset of information. For example, an individual might be able to specifiy that the data of interest lies between T + 20 hours and T + 22 hours (T representing launch time). This type of subdivision is dependent on the manner in which the data is recorded.

Because of the uncertainties of the data formats, etc., it is not possible at this time to speculate on any general formats which could be recommended as being most useful or advantageous. It is recommended that further study be undertaken to determine if any set of general formats are feasible. This is of interest principally to provide some standardization or homogeneity to the AAPA data base.

It is not considered wise to attempt to provide a reformatting service to users as it would be virtually impossible to anticipate all possible formats. In fact, this type of service would probably have little real value to a user. The user must design his computer program to accomplish his intended manipulation of the data and his acceptance of the data "as is" from AAPA would have little impact on the cost or complexity of his program generation. In essence, AAPA should only provide identical copies of information in the data base along with an appropriate verbal description of the particular data format

The AAPA should, however, attempt to produce tapes which are compatible with the computer system available to the user. The most widely acceptable tape characteristics are considered to the "IBM compatible". This is to say that AAPA should have computer facilities available to it to produce 7 or 9 level tapes in densities of 200, 556, 800, and 1600 bits per inch. Furthermore, a program capability should exist to adjust blocking factors so that the number of bits, as a record, could be adjusted in a reproduction process to accommodate the users requirements. Finally, it is recommended that all tapes be recorded in the "binary" mode as opposed to character grouping. Character compatability between various computer manufacturers leaves much to be desired and can create a multitude of problems.

3.10 Distribution Scheme

The method chosen to accommodate the handling of user queries and the distribution of outgoing data will depend upon the structure of the archives whether centralized or decentralized, the volume of data being distributed, and the attitudes of the users - particularly toward cost effectiveness. In a decentralized structure, it is anticipated that the decentralization would be accomplished on a topical basis (See Section III) if at all; i.e., some of the data will be stored in existing facilities according to the subject content of the data. These facilities, which may be controlled by NASA or by other government agencies, already have dissemination channels which are perfectly acceptable. For example, the Scientific and Technical Information Division of NASA has one of the most extensive data dissemination networks of any agency in the country; the Defense Documentation Center and the Technical Information Center of the American Institute of

Aeronautics and Astronautics both have very good distribution channels; and both the National Library of Agriculture and the Library of Medicine, which support research organizations within the federal government, have established distribution channels which could serve to disseminate experiments data.

If the AAPA develops as a self-contained entity, it can have its own distribution mechanism which will evolve with the system. At the outset, there will be a relatively small data base and few user requests. It will be necessary, however, to emphasize service to the users if there is to be a transference of AAP data to users and if users are to develop any confidence in the value of the AAPA. The user request and data dissemination could be handled within the archive itself, but the door must be left open in the design for growth and expansion. As more AAP missions are flown, the data base broadens to interest more users, and the publicity of the AAPA becomes more widespread, the point will be reached where it becomes more practical and efficient to create an independent, yet internal, branch to handle user requests and data distribution. Again, this branch would be a functional part of the archives, but established solely to interface withthe users. A third existing possibility is to create some independent, monolithic input/output agency to service all federally operated data centers in correlating user data requests. A similar system has been proposed by MSC to serve NASA in distributing pictorial data; i.e., that a commercial organization should actually be the distributor of (at least) all NASA generaged pictorial data. This concept might have practical value, especially with interagency operation where the scope of data becomes so broad.

As will be pointed out in Section 4.0, it is believed that at least initially, the AAPA should be a self-contained system storing all non-space oriented AAP data (the space oriented data to be stored within NASA at the NSSDC). As such it will have its own distribution scheme as outlined above. Since the system will service government, industry, and the scholastic community, the detailed nature of the distribution mechanism will accordingly depend both upon the users' facilities and the frequency with which each user queries the archive. Because the system is dealing to a large extent with digital data (probably stored on magnetic tape), a direct wide band, high speed data link system which connects the AAPA with individual users is feasible for the more frequent users. The wide band data link would have the advantages of facilitating the distribution of large blocks of data, providing the user direct access to a digital AAPA catalog, and of enabling the user's own computers to restructure the data as they are received to accommodate his own particular formats and programs. For other users who may require less frequent access or who cannot afford such a sophisticated system as the wide band data link, distribution may be accomplished by mail or some similar means. Finally, for users who require many photographs, facsimile transmission systems will permit review of low resolution copies on the user's premesis without his traveling to the "archive". For the pilot system distribution should be solely by mail until user need and interest dictates otherwise.

3.11 Initial System

Because of the nature of the Apollo Applications Program, the data to be stored in the archive will be accumulated in steps over a period of several years. Thus, the archives itself will gradually evolve from a small pilot system, less complicated than that discussed previously.

3.11.1 Scope. - The initial system, aimed at fulfilling the system requirements through a limited time period (about two to three years), will provide a basic working system, conservatively outfitted but capable of multidirectional expansion. While the system is evolving, additional user information can be gathered, and as this data accrues, the systematic expansion can then be planned to enable a closer fit between system demands and performance capabilities. This approach is in keeping with the general economics of a system such as the AAPA. Since the utilization requirements will be small at first and grow with time, there is no justification for implementing a full-blown system at the outset when it would not be utilized.

While the pilot system will be smaller and less complex than the full-blown system, it must still fulfill all of the system functions outlined in Figure 15, namely, the active collection, organization, announcement, storage, retrieval and dissemination of data. The emphasis in the beginning, however, will be placed on organization - the development of vocabularies, surrogation techniques, paperwork procedures, etc. and on dissemination and promotion. A possible configuration for the pilot system, shown in Figure 16 consists of a shipping and receiving group, technical writers, a use analysis and indexing group, some data processing personnel, and a staff group.

The staff group will be responsible for the management of the archives and will coordinate all activities with the experimenters, the users, other data facilities and NASA field offices. In working with an experimenter, the group will establish which data is to be supplied and will monitor the progress of his analysis to insure that a delivery schedule for the data is adhered to. The tasks of monitoring the progress and of acquiring the data will be relegated to acquisition specialists who have an appropriate mix of technical background. After the data has been received and processed, the group will handle user requests and maintain statistics of data flow for advanced system planning. Thus, in terms of the essential system functions, the staff group will be responsible for both theactive collection and the retrieval of data in addition to the overall system management.

The use analysis and indexing group will be responsible for the task of data organization. The group must have personnel capable of discerning the needs of the users such that the best tags may be applied to the experiments data for user access. These data use specialists and data control analysts are in extremely short supply, and the shortage is not expected to end in the forseeable future. Obtaining capable personnel to organize the pilot system could therefore pose serious problems.

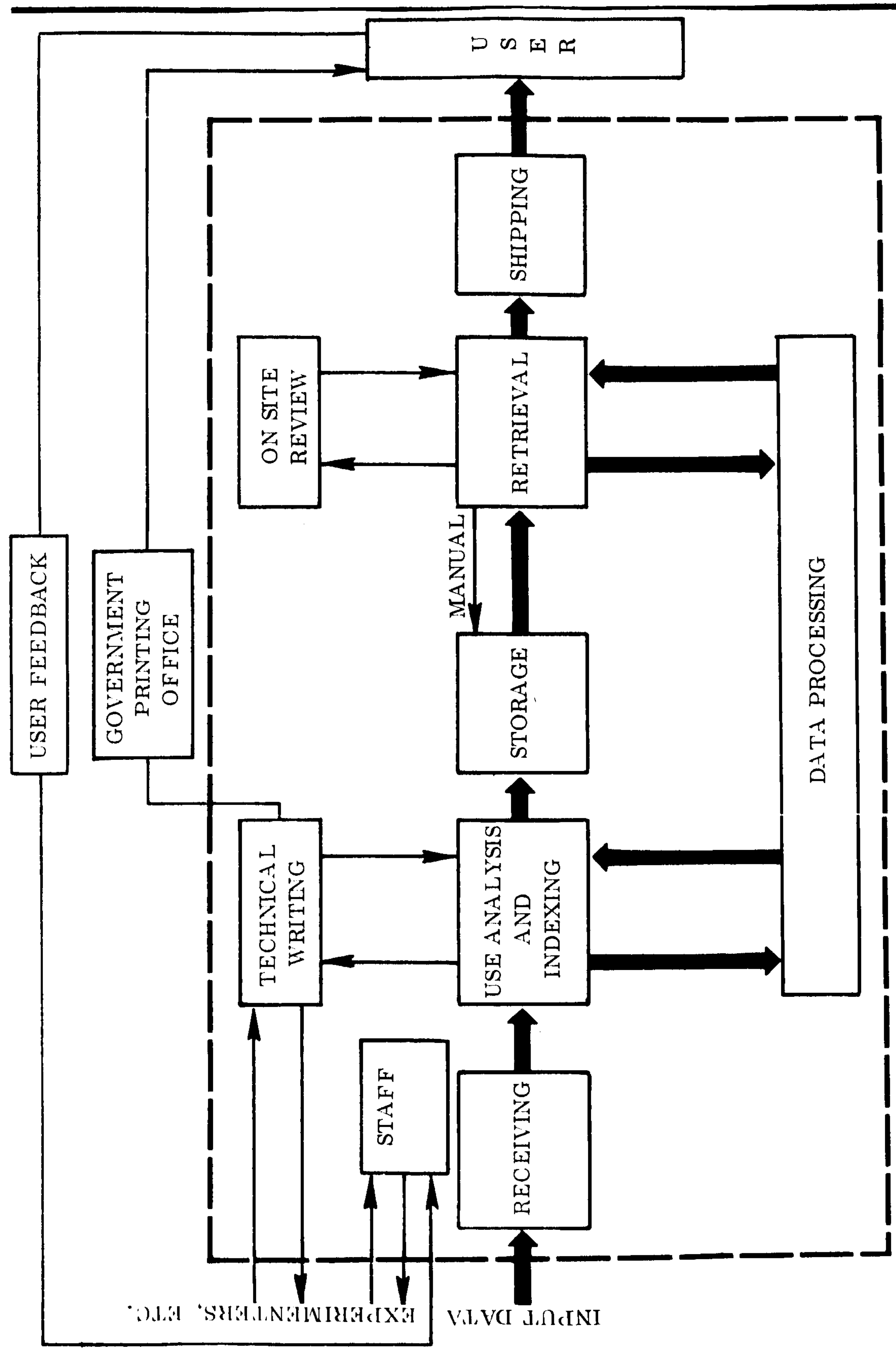


Figure 15 Pilot System Concept

Technical writers will work closely with the experimenters and the indexing personnel to prepare catalogs and briefs on all experiments data stored in the pilot system. The briefs, rather than merely listing the data which is being stored in thearchives, should be designed to assist potential data users to decide if they can make use of the available data. Such things as a description of the experiment and the measurements taken, the background of the experiment and the experimenter, the data reduction techniques employed, the format of the available data and a list of references should be included.

The amount of data processing required in the pilot system is expected to be rather limited, though some will be required. That which will be done will set the pattern for the "full blown" system.

It has been suggested that the initial system should not be staffed by NASA personnel but rather by subcontract personnel. The subcontract task would be to coordinate between NASA and industry or academic users. The operation of the Archives would require interfacing with all NASA centers and related government agencies for the input data and information which would be identified, analyzed, procured, coordinated, controlled, and distributed by this group. The output of this group would be experimental data and information in standard formats, available upon request from the respective NASA centers. The advantages of subcontracting this task would include 1) accessibility to needed disciplines without robbing existing government facilities; 2) evading any direct contact with industry or academic institutions at delivery time; 3) providing a bufferi ng organization between a product oriented group and scientific research; 4) preventing any legal exposure to NASA as to the proprietary rights of the user. Several disadvantages in subcontracting this task which exist include the antithesis of items 2 and 3 above and a) training of additional disciplines within government organizations to transfer to operational systems at a later data b) a direct interface with the user would be desirable in order to create practicality from this experimental data being generated.

3.11.2 Personnel Considerations. - As mentioned above, one of the key problems facing the pilot AAPA is that of acquiring capable personnel. The supply of these personnel is extremely short, and there have been a number of instances of severe operational disturbances in major systems caused purely by a mismatch between the personnel requirements of the given systems and the actual manpower supply. The training of new or additional system staff in the specifics of their responsibilities within the system's detailed mode of operation has been a serious and continuing problem for most information systems of any size. Rarely assessed adequately in advance by system designers the training problem can cause serious cost and productivity inefficiencies and must be taken into account in the design process.

Professor A. J. McNair of Cornell University further emphasized the personnel problem in a discussion with UAC personnel. He feels that the shortage of capable and qualified professionals in the fields of photogrammetry, photointerpretation, and geodesy will create a very serious problem in the future practical uses and applications of space photographs, and substantiated this position with some interesting statistics. At the present time the United States presently averages only one graduate per year in these fields for every 5,000,000 people in our population. Europe and Canada graduate on the average of one man per 100,000 people per year.

At the present time, Austria is the only country in the world of which Professor McNair knows that has ever had a surplus of professionals in photogrammetry, geodesy, or photointerpretation. Not only does this predict dire consequences for the future, but there are definite signs that the present situation is very grave. For instance, the Army Map Service, in order to get a man with sufficient experience and capability to head up their research and top level scientific projects was forced to import Professor E. Aene Bjerhammer from Sweden. There are only 15 universities in the United States which offer cirricula in the fields of photogrammetry, geodesy and photointerpretation. Professor McNair pointed out that survey and civil engineering firms subscribe to the following ratio regarding personnel they employ. For each professional they would like to have 3 technicians and approximately 7 unskilled laborers. The problem at the present time and even more in the future is not in finding technicians and unskilled laborers but rather in finding the professionals. As Professor McNair pointed out earlier, the lack of professionals might be our most serious problem in the use of our space photographs.

3.11.3 Budget Specifications. -

3.11.3.1 Introduction. - Minimal budget specifications for the initial years of operation of the AAPA system follow. A discussion of functional and line details appears as section 3.11.3.2 below and tables of line and function items as section 3.11.3.3.

These details, however, need to be understood inthe context of a number of qualifications. The budgets cover a three-year period but re-estimates of the second year and later costs will be required subsequently as feedback from initial operations and design iterations becomes available. The budgets cover only direct-type cost and include no indirect overhead, floorspace, fringe benefit or other indirect costs such as purchasing, contract administration, hiring costs, etc. Minimal allocations for data record rewriting and reproduction have been made on the assumption that these activities are relatively minor in the first year of operation. The allocations for printing are intended for use promotion materials, the AAPA Index and AAPA Catalog and similar items and should not be confused with data record reproduction. The budgets are presented in two different forms for the sake of clarity. A line budget follows the usual salary-by-salary, item-by-item organization. A functional budget distributes all of these costs over the matrix of six major functions which the AAPA will have to accomplish in carrying out its task: Data Acquisition, Surrogation and Control, Storage and Retrieval, Disseminationand Promotion, Administration and System Development. The dollar totals will be found to compare in rather general terms with the initial costs of other information systems recently developed such as the Air Force Machinability Data Center. No program budget is included. Development of a program budget before more detailed specification of operations, output products and the priority allocations among users would be premature. The budgets presume, of course, a reasonable (9 months) allocation of get-ready time before full operation.

The budget estimates are actually quite low considering the cost of the Apollo Applications Program and the considerable political pressure for secondary uses of the AAPA data. The minimal size of the initial budgets may be attributed to a number of causes:

1) Commitment to a large-scale system has been deliberately postponed.

2) Field offices have been avoided by the use of centrally located industry specialists with adequate travel funds.

3) Data archiving outside the AAPA store has been designed into the system not merely tolerated.

4) Use of tertiary data created outside the system has been provided by careful management of interfaces.

5) A selective approach to the promotion of data use is adopted rather than a shotgun approach.

6) Some output products and services use the existing information networks and computer programs of other agencies.

3.11.3.2 Discussion of Detail. - The initiation of the system will be marked by high start-up costs for system development which are recognized openly by the functional budget. During the first year in addition to the time of the systems man and considerable machine time, the functional budget allows considerable participation by the data analysts and others in the development of vocabularies, surrogation techniques, paperwork procedures and the like. A small amount of cost for external systems work is also allowed principally to pay for the cost of adaptations of the output programs of other agencies whose systems are being shared. An example of such an adaptation might be the revisions of index-printing programs made for NASA's Reliability Abstracts and Technical Reviews from the original DOCINC programs created for STAR. In the third year, $40,000 for systems development is allocated for a formal internal or external appraisal of progress and redesign of the system.

Actual control, surrogation and storage costs are relatively low at first partly because little true AAP archival data is coming in. During this period, much of the time of the staff who would normally work at these functions is being committed to system development work. Administration costs are also kept relatively low throughout the program because of full recognition of system development problems elsewhere in the budget and in contracted development efforts previous to it. Dissemination and promotion costs (function D) are shown as relatively significant even in the first year-- 24 percent of the total-- because the data use specialists will be involved at an early date in the actual task of finding users for data not yet available and because printing costs for function D reflect the dissemination of early types of promotional literature.

Rises in salary costs during the three-year period are chiefly caused by the addition, in the second year of operation, of an executive officer (administrative assistant) to assist the administrator and guarantee him adequate time for nonpaperwork activities. A second clerk in the control activity is also added inthe second year of operation.

Some detail on line items is also desirable at this point.

The administrator is staffed at level approximately equivalent to full professor, and the possession of a doctorate (probably in geophysics, surficial geology, or an appropriate branch of economic geography) is required in order to ensure a proper appreciation of the opportunities in the data as well as a high estimation of the archive by the people with which it must deal. A master's degree and extensive experience in exploration geophysics will be an acceptable but a less satisfactory substitute. Extensive experience in information systems work is not required.

The activities and qualifications of the Data Acquisition Officers are discussed elsewhere inthis report. Their chief expertise lies in coping with life amid the bureaucracy particularly the NASA and scientific establishments.

The data control analysts are a mix of personnel involved in the descriptive analysis of data records and in relatively low level analysis of potential uses. A mix of backgrounds for this staff will be not only desirable but necessary. A number of mixes are possible. One such mix would be a photoanalyst, a person with equivalent experience with magnetic tape records and a person whose experience lies in the creation and manipulation of the kind of records used in libraries, document control agencies such as DDC or project-management systems. The data analysts are supported by clerical personnel and machine routines, and the recording of data generated by them is handled onsite via a paper-tape punching typewriter.

The activities of the systems man and data use specialsits have been discussed elsewhere in this report.

3.11.3.3 Budget Tables

AAPA FUNCTIONAL BUDGET COSTS
(THOUSANDS OF $)

GENERAL

	PROGRAM AREA	YEAR 1	YEAR 2	YEAR 3	TOTAL 3 YEARS
A.	Data Acquisition	22	50	50	122
B.	Data Surrogation and Control	33	52	52	137
C.	Data Storage and Retrieval	28	33	48	109
D.	Data Dissemination and Promotion	77	121	127	325
E.	Administration	29	41	45	115
F.	System Development	114	23	72	209
	TOTALS	303	320	394	1,017

NOTES: Indirect overhead, floorspace, and certain other costs are not included. See narrative, elsewhere, for discussion of details.

Detailed Breakdown

AAPA Function Line Budget Costs (Thousands of $)

		Function Areas - Expenditures																	
		Year 1						Year 2						Year					
	Items	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
SALARY LINES	Administrator					18						20						22	
	Secretary					5						5						5	
	Executive Officier											10						12	
	Data Acquisitions Officers	15					30	45						45					
	Data Control Analysts		15				15		30						30				
	Data Control Clerks		3				2		10						10				
	Data Use Specialist				30		30				60						50		10
	Data Use Clerk			2	3					2	3					2	3		
	Software/Procedures Specialist						15						15						15
ITEM LINES	Office Equipment	2	8		2	2	1		2		1	2			2			2	1
	Archival Storage Equipment and Supplies			25	5					30	10					45	15		
	Office Supplies	2	2	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2	1
	Printing				20						30						40		
	Processing Machine Time		5		5		10		8		5		2		8		7		5
	External Systems Work						10						5						40
	Travel and Maintenance	3			10	2		3			10	2		3			10	2	
	TOTALS																		
	A Data Acquisition	22						50						50					
	B Surrogation and Control		33						52						52				
	C Storage and Retrieval			28						33						48			
	D Dissemination and Promotion				77						121						127		
	E Administration					29						41						45	
	F System Development						114						23						72

NOTES: Indirect overhead, floorspace, and certain other costs are not included. See narrative, elsewhere, for discussion of details.

AAPA LINE BUDGET COSTS
(Thousands of $)

SALARY DETAIL

	Line Item	Year 1		Year 2		Year 3	
1.	Administrator	(1)	18	(1)	20	(1)	22
2.	Executive Officer			(1)	10	(1)	12
3.	Secretary	(1)	5	(1)	5	(1)	5
4.	Data Acquisitions Officers	(3)	45	(3)	45	(3)	45
5.	Data Control Analysts	(3)	30	(3)	30	(3)	30
6.	Data Control Clerks	(1)	5	(2)	10	(2)	10
7.	Data Use Specialist	(4)	60	(4)	60	(4)	60
8.	Data Use Celrk	(1)	5	(1)	5	(1)	5
9.	Software/Procedures Specialist	(1)	15	(1)	15	(1)	15
	TOTAL	(15)	183	(17)	200	(17)	204

NOTES: Figure at right of each column is dollar cost; figure at left in each column, in parentehses, is the number of such positions.

AAPA LINE BUDGET COSTS
(Thousands of $)

NONSALARY DETAIL

	Line Items	Year 1	Year 2	Year 3
10.	Office Equipment and Furnishings	15	5	5
11.	Archival Stroage Equipment and Supplies	30	40	60
12.	Office Supplies and Postage, Telecommunication	10	10	10
13.	Printing	20	30	40
14.	Processing Machine Time	20	15	20
15.	External Systems Work	10	5	40
16.	Travel and Maintenance	15	15	15
	TOTAL	120	120	190

4.0 DEGREE OF DECENTRALIZATION

By its very nature, the total scope of the AAP experiment-related data is extremely widespread. Topically, there are experiments which are oriented toward Earth and Lunar resources, biomedicine, meteorology, communications, the space sciences, biology, and space systems engineering. The responsibility for the experiments can be held at many different geographical locations within the NASA organization. Principal investigators are located throughout the world. Subsequent users of experimental data will not only be interested in each topic but will be at many geographical locations. Thus, in order to decide upon a configuration in terms of centralization or decentralization of the data archives and upon the degree of potential redundancy required, it is necessary to examine and weigh the topical and geographical considerations. Other factors for consideration are the mechanics of data acquisition, the sources of input data, and the requirements of the users.

The topical definitions have not been made in concert with the rigorous academic or library definition, but more oriented toward economic discipline definitions. These definitions, as given, relate more readily to the various sectors of the economy, and in some cases, there is inherently inferred a geographical relationship. For example, Earth resources may be divided into at least the two subsectors of minerology and agriculture. The principal interest in one form of minerology exists in the western United States, while one form of agriculture interest is in the middle west. In other cases, such as biomedicine or communications, this relationship cannot be seen as clearly. Even the subsectors are not geographically relatable because of the extreme diversification of this type of industry. Thus, to centralize or decentralize on the basis of area of geographical preference for a given type of data usage becomes a very complex problem.

4.1 Geographical Decentralization

The analysis for geographical decentralization is dependent upon the sources of input data, the acquisition of the input data, the user locations and requirements, and interest in a given topic on the one hand, and the cost of duplicate equipment and storage, as well as a high speed distribution system, on the other. The primary benefits of geographical decentralization are a reduction in acquisition costs and convenience for the users. The disadvantages are the costs of duplicate storage and support activities.

Data received at the AAPA will come from essentially three sources; viz., NASA field offices responsible for experiments, the experimenters themselves, and some secondary input sources, such as other related data centers or academic institutions. Since the hardware and support facilities developed for the Apollo project are to be used to the maximum extent possible in the Apollo Applications Program, it is logical to assume that the postflight data flow for the AAP will be very similar to that of Apollo. This data flow, discussed in detail in Section 3.7.3, indicates that original tapes of all raw data will be routed to Houston where they will remain (for approximately a year), until being placed in archival storage at Goddard. Thus, selected raw data from the AAP experiments should be supplied to the AAPA from either Houston or Goddard, depending upon the time factor involved.

Similar to the procedure employed by the NSSDC, it is expected that the experimenter will be under a contractual obligation to NASA to furnish his reduced and final analyzed data to the AAPA. The reduced data refers to the raw data which the experimenter has calibrated, correlated, processed to remove noise and put in some useful form as a function of time, distance, et cetera. The final analyzed data refers to that data which has undergone further analysis, and is indicated by the experimenter as being most representative of his findings. In the process of obtaining this data, it may become necessary for AAPA representatives to visit the experimenter, or at least travel to the NASA field office responsible for the experiment. In the event that experimenter visitations are required, this should be accomplished with the cognizance and approval of the responsible NASA Program Manager.

The acquisition of data for the AAPA will certainly require a centralized control function. This control will be necessary for economic reasons, whether the distribution of data is centralized or decentralized. The nature of requests for experimental data will, in all likelihood, require the delivery of intraexperimental data which may have been stored or procured at different NASA centers, or by different Principal Investigators. The Data Acquisition Group of the AAPA serves the function of coordinating with the input sources of primary, secondary, tertiary and quarternarv data.

The sources of experimental data can be defined as primary and secondary. Primary sources are NASA groups and experimenters working directly on the Apollo Applications Program. Secondary sources are the support groups in NASA and other Government agencies, such as the National Weather Record Center and U.S. Department of Commerce. If the data is to be distributed by AAPA, these primary and secondary sources of experimental data and information will interface with the AAP through the archives system. A centralized coordination group within the archives system will prevent the redundant acquisition of experimental data.

The requirements of the users of the experimental data do not, at this time, justify a decentralized system of distribution. The types and forms of experimental data that will be available for distribution, according to the present experiment listing, do not require real time or near real time distribution. The number and location of potential users presently interested in the experimental data do require a source of information concerning AAP experiments, but during the initial phase of the program, the demand for experimental data is unlikely to be so great as to justify the increased facilities. If an increase in the number of users is to be desired, then a centralized system of distribution will better accomplish the dissemination of information and publicity that will be required. Consideration must be given to the initial design to allow for this increase in experimental data usage. An increase in the number of users, and a requirement for near real time acquisition of data, will augment the reasons for decentralization. These considerations should include the possibility that this archive will be a model and training ground for an operational satellite system with an associated archive and distribution system.

As mentioned previously, one of the functions of the Data Coordinating Group will be to direct users to other data centers which are storing pertinent data not stored at the AAPA. Should the AAPA receive enough requests for the same block or type of data which they are not storing, it may then be desirable to obtain copies of such data from the secondary sources, and store them in the AAPA. These secondary sources, some of which are listed in Section 3.7.6, are scattered throughout the country.

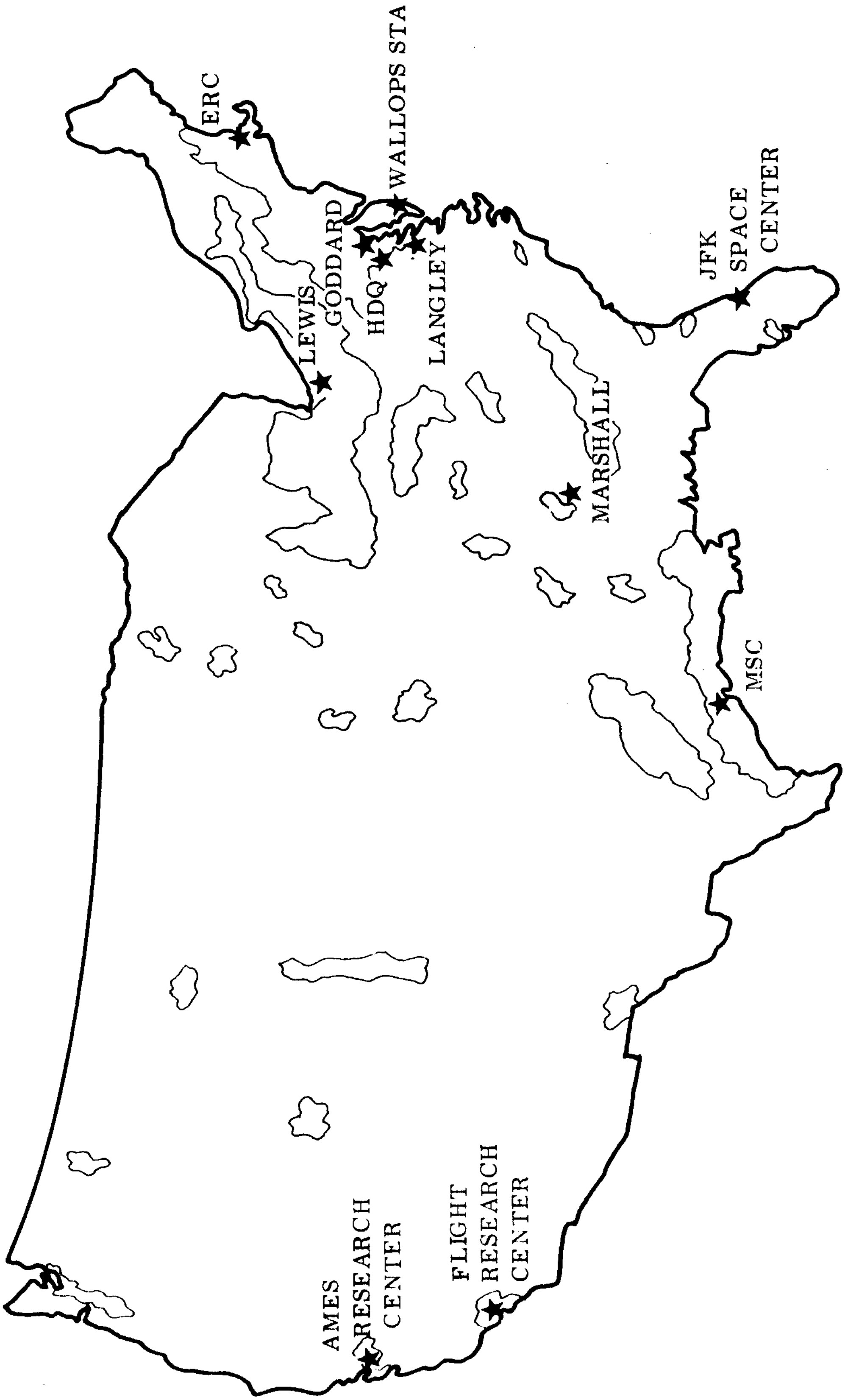
The geographical relationships of inputs and output were reviewed during the course of the study. Shown in Figure 16 is a map of the United States with the major U.S. industrial centers. There was not sufficient time nor money to relate the Standard Industrial Classification Numbers (S.I.C.) against these industrial areas during this study, and this is recommended to be accomplished during a folow-on work scope. The S.I.C. Nos. were assigned to the various experiments and specific companies. This relationship was used as a basis for determining industries which should be briefed about AAP. The next logical step would be to identify each of the industrial areas by S.I.C. No., and relate each area to potential experimental interest. It is believed that additional information is readily available in currently published data from the U.S. Dept. of Commerce. However, Figure 17 (Geographical Review of Users) when overlaid on the map identifies the potential users who were contacted during the study. These potential users were located in many of the areas identified on the map. A representative industry should be contacted from each of the remaining areas in order to complete the geographical area review of users.

Figure 18 (NASA Centers) shows the relationship of the NASA centers to the industrial areas. These NASA Centers can be considered to be the primary sources of experimental data for AAPA, even though the experimenters may be located throughout the world. Because the experimental data is sent to the experimenter from one of these locations, in order to reduce the support requirements, it is suggested that these centers furnish duplicate copies of the experimenters information to the AAPA. This copied experimental data may be retained at the center as opposed to physical shipment of the data record, but notification of its availability should be made available to the AAPA.

Those centers visited during the course of the study are shown in overlay Figure 19. This figure is of particular significance because these centers are the ones which are most involved in the AAP data flow.

68-17

Figure 16 Major U.S. Industrial Centers



68-16

Figure 15 NASA Centers and Industrial Areas

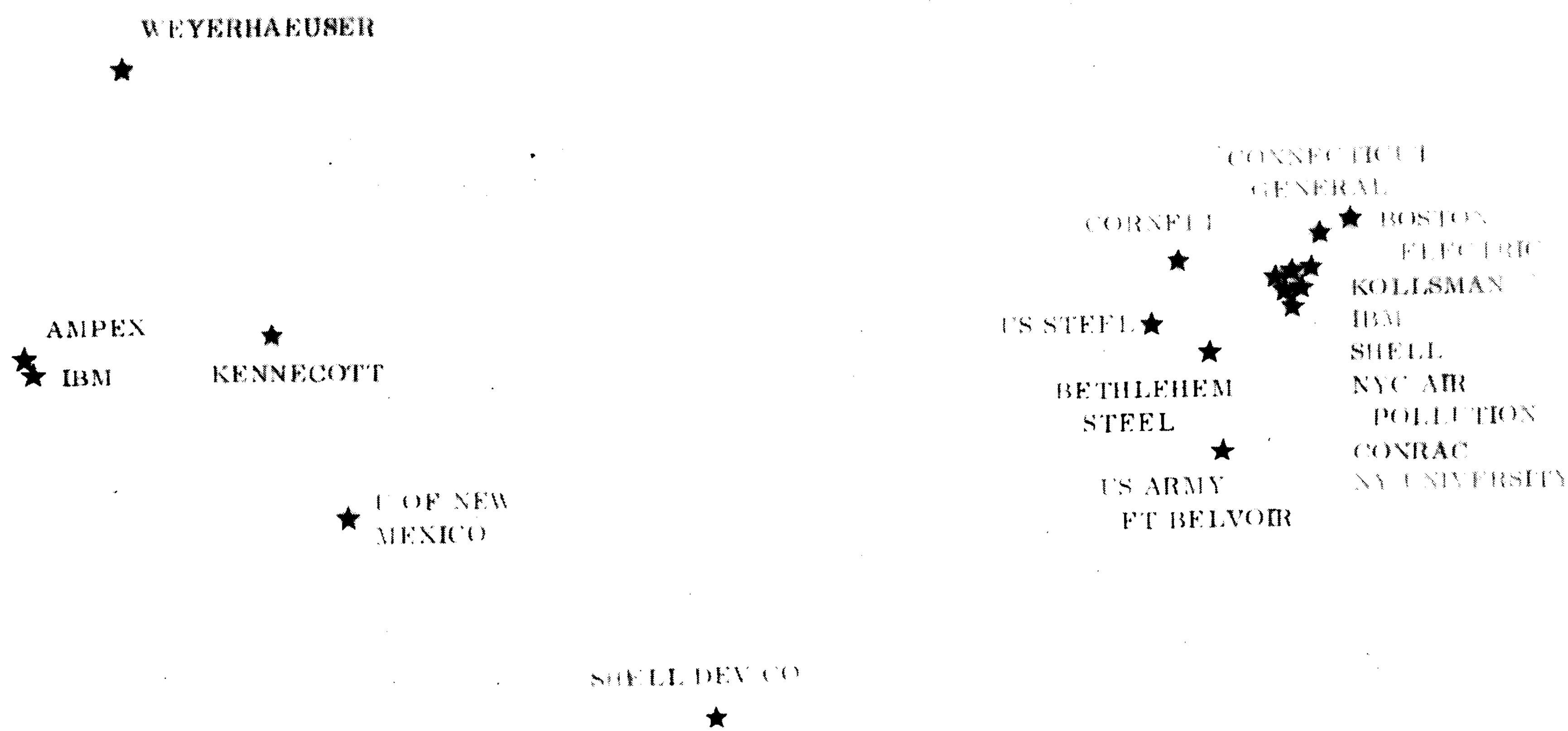


Figure 17 Geographical Review of Users

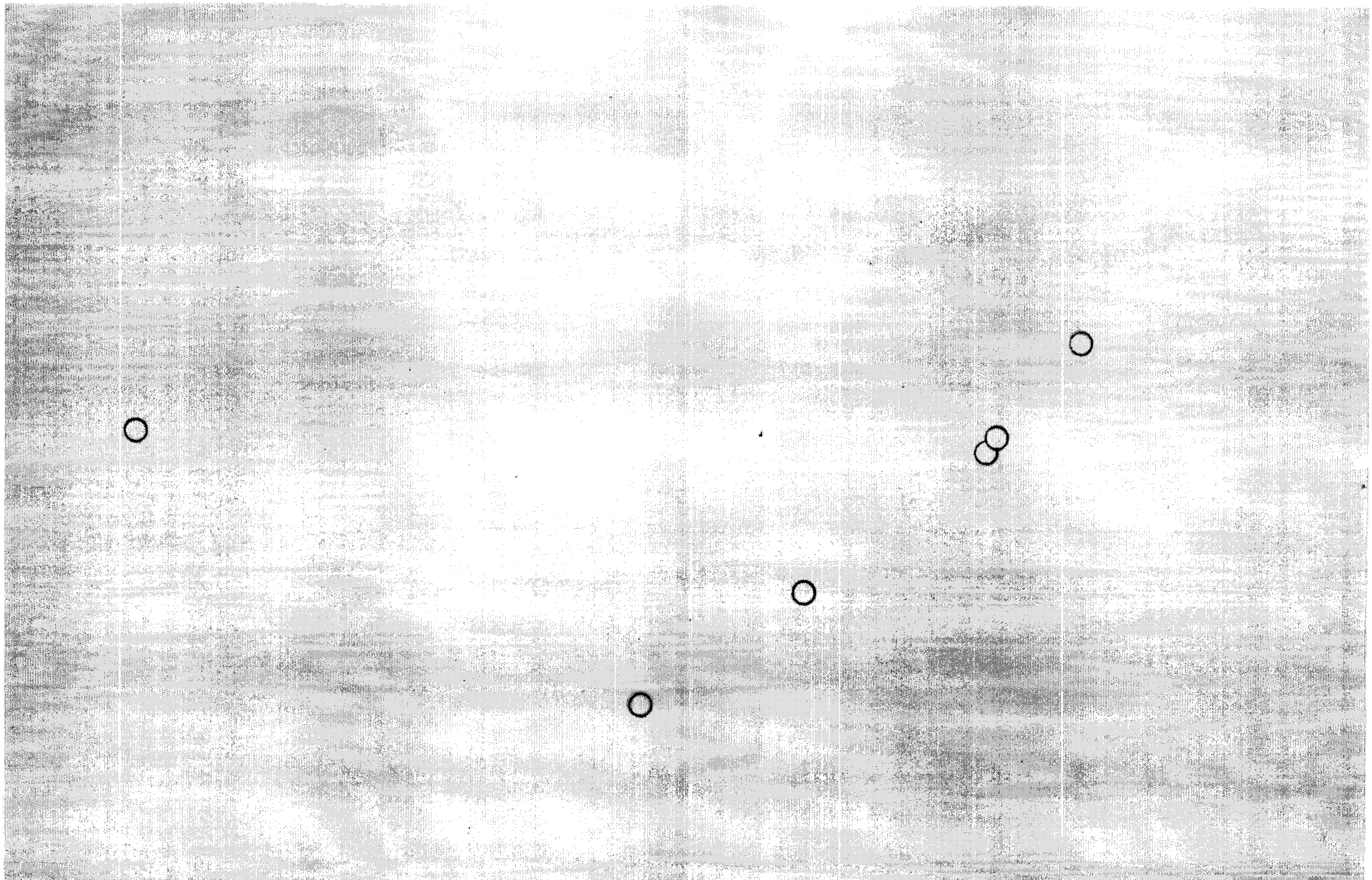


Figure 19 NASA Centers Visited

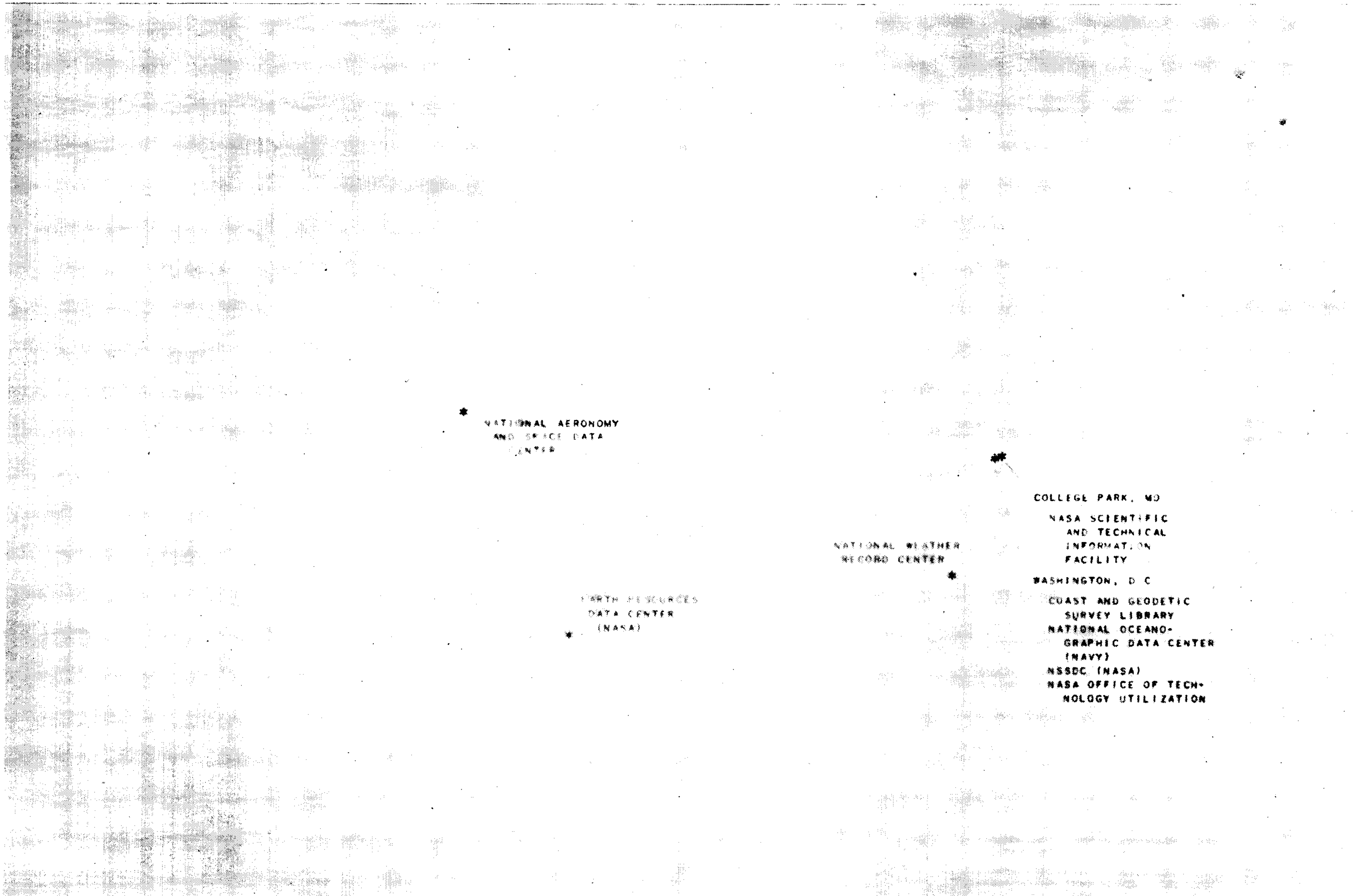


Figure 20 Secondary Sources

The secondary sources of information and experimental data are shown in overlay Figure 20. Liaison with these secondary sources must be maintained by the AAPA, in order to fulfill supplementary data needs. Many of these sources will also be users of data within the federal structure.

4.2 Sources of Input

As defined in the preceeding section, the primary sources of input are those groups working directly with the AAP, and the secondary sources are support groups in NASA and other government agencies. Figure 21 (Experiments Data Input) relates the data types that will be available from these sources for the AAPA. The classification of data is also shown by the symbols. The number of data types that will be available suggests that very close coordination must be maintained between the input sources and the AAPA and hence fortifies the argument for a centralized AAPA to prevent redundancies of effort.

Figure 22 (Experiments Data Function Input) shows a different relationship between the primary sources and the function which these sources might serve in AAP. Certain users will require the functional data described here in conjunction with the experimental data types depicted in Figure 21. In this case, the function data would be used as supplementary or definitive information which might be used for purposes other than those for which the original experiment was designed. A data package under these conditions would require a higher degree of coordination for assemblage and thus fortifies the argument for a centralized AAPA.

The input to the AAPA from secondary sources is shown in Figure 23. The data types that can be expected to be required are plotted against some of the secondary sources in a manner to indicate the classifications expected from these sources. The data functions are not shown here because the number of functions this data would fulfill is as varied as the potential user's problems. Although the secondary source may appear to have the effect of decentralizing the AAPA distribution system, again a centralized coordination group should be used to provide the potential user with a complete work package.

4.3 Distribution Considerations

A distribution system must not only consider the demand for the product which is to be distributed but the mechanics of distribution with respect to the existing distribution facilities. A centralized distribution system may fulfill the requirements of throughput, turnaround time and compatability with the existing distribution facility better than a decentralized distribution system when demand for the product has not been defined in term of quality and quantity.

Data Types

Primary Sources	Digital	Analog	Photographic	Hard Copy	Procedural	Space Craft Characteristics	Initiating Implementation
NASA Headquarters				Δ	Δ		
Goddard	O X	O X		O X			
Langley						Δ O	
ERC							
Wallops Island						Δ O	
Lewis						Δ O	
Marshall (MSFC)	O X	O X	O X	O X		O	
JFK							Δ O
Houston (MS)	O X	O X	O X	O X			
Flight Research Center						O	
Ames Research Center						O	
Experimenters	X O	X O		X O Δ			

Classification

X - Primary
O - Secondary
Δ - Tertiary

68 27

Figure 21 Experiments Data Input

Data Function

Primary Sources	Ephemeris	Sensor (RAW)	Space	Equipment	Environmental	Experimental
NASA Headquarters		Δ	Δ	Δ	Δ	O Δ
Goddard	O X		X O			
Langley				X O		
ERC				X O		
Wallops Island						O X
Lewis				X O		
Marshall		X O	X O	X O	X O	O X
JFK				X O	X O	
Houston (MS)		X O		X O Δ	X O	X O
Flight Research Center						O X
Ames Research Center						O X
Experimenters						X O Δ

Classification

X Primary
O - Secondary
Δ - Tertiary

Figure 22 Experiments Data Function Input

Secondary Sources	Digital	Analog	Photographic	Hard Copy
	Data Type			
ESSA	X O		X O	X O
National Aeron	△		△	△
Coast and Geodetic Survey Library			X O	X O
Nat'l Oceanographic Data Center	X O △	X O △	X O △	X O △
Nat'l Weather Record Center	X O △		X O △	X O △
NSSDC	X O △		X O △	X O △
Earth Resources Data Center	X O △	X O △	X O △	O △
NASA Scientific & Tech. Info. Fac.	O △			O △
NASA Office of Tech. Utilization	O △			O △
U. S. Dept. of Agriculture	X O		X O	X O
U. S. Dept. of Commerce	X O			X O
U. S. Dept. of Transportation	X O			X O
U. S. Dept. of Defense	X O	X O	X O	X O
U. S. Dept. H. E. W.	X O			X O

Classification

X - Primary
O - Secondary
△ - Tertiary

68 29

Figure 23 Secondary Sources Input

The througput capabilities of a centralized system of distribution becomes unwieldy when the number of requests for various items of experimental data reaches the level of interference of requests. Because of the nature of AAP, it is not anticipated that this level will be reached in the early phase of the program. The coordination required to secure deliverable experimental data suggests that centralized control be maintained. A decentralized distribution to the user could be made to operate in parallel with this control but it appears that the throughput capabilities would not be increased to the extent that it would warrant the increased cost and support effort required.

Consideration has been given-to the output product. The Output Experimental Data Types are compared with the topical sectors of industry in Figure 24. The classification of the data types are shown related to each sector by the appropriate symbol. This estimate is based upon the experience gained during this study where direct contact was made with representative industries. Since this chart represents experimental data only, it does not reflect the auxiliary or supplementary data that may be required from secondary sources. However, in some cases it appears that the supplementary data may go direct from the secondary source to the user and therefore not be a contributing factor to the throughput capabilities or existing distribution facilities.

Figure 25 is a representation of the output package of experimental data by classification. This chart reflects the experience gained during the actual contacts with the organizations contacted during this study. Since these contacts were rather brief the assignments shown here are initial estimates and will change as a given organization becomes more involved with AAP experiments data. However, the indications are sufficient enough to predict that the various data types depicted in Figure 25 will be needed within each sector.

The turnaround time for the AAPA is defined as being from the point of initial request until delivery of an experimental data work package into the hands of the potential user. The processing of this request, locating the experimental data, inspecting the data, processing, packaging and delivery is included in the turnaround time.

Of those potential users contacted, initial estimates were made that a turnaround time of one week to three weeks would be desirable. These estimates were colored in one way or another depending upon specific problems such as user inspection of the data, degree of definition, and whether or not the relationship between the company and AAPA was continuous or sporadic.

In the initial phases of AAPA, it will be more practical to obtain this type of turnaround time in a centralized system than noncentralized. The functioning and

Data Types

USERS	Digital	Analog	Photographic	Hard Copy
Resource Oriented	X O Δ	X O Δ	X	X O Δ
(Insurance) Services & Misc.	O Δ		Δ	O Δ
Non-metal Product Mfg.	X O	X O	X O	X O Δ
Agriculture	X O	X O	X O	X O Δ
Metal Product Mfg.	X O	X O		Δ
Optics & Instrumentation	X O	X O	X O	Δ
Communications	X O Δ	X O Δ		Δ
Mechanically Oriented	X O Δ	X O Δ		X O Δ
Research & Development	X O Δ	X O Δ	X O Δ	X O Δ

Classification

X - Primary
O - Secondary
Δ - Tertiary

68-30

Figure 24 Output Experimental Data Types

Organizations Contacted	Primary	Secondary	Tertiary	Quarternary
Boston Edison			X	X
Cornell	X	X	X	
Carviac				X
N. Y. C. Air Pollution	X	X	X	
U. S. Army	X	X	X	X
Univ. Of New Mexico	X	X	X	X
Conn. Gen. Life Ins.			X	X
New York University			X	X
Kollsman Inst.	X	X	X	
Weyerhauser	X	X		
Kennecott Copper	X	X	X	X
IBM	X	X	X	X
Ampex			X	X
NASA/Ames		X	X	X
Shell Oil	X	X	X	
Shell Develope	X	X	X	
Bethlehem Steel	X	X	X	
U. S. Steel	X	X	X	X

68-31

Figure 25 Output of Experimental Data by Classification (Estimated) coordination of tiers of organization would tend to increase this turnaround time and result in poor service to potential users.

Consideration was given to the existing distribution facilities during the course of this study. Natural divisions that occur are the organization structure and the type of delivery service. The organization structures may be those of the Office of Technology Utilization, the secondary sources, specified centers which hold experiment responsibility, or existing library facilities at Goddard Space Flight Center. The delivery service may be direct wire, airplane, U. S. Mail or hand carried.

Although it was recognized that none of the existing organization structures were specifically program oriented towards the AAP and that none had the complete capability for handling experimental data, it was thought that the distribution of experimental data from the AAPA might be centrally distributed through one of these organizations. A preliminary investigation reveals that for one reason or another this form of centralized distribution is impractical at this time. Yet, each of these organizations acts as a distributer of topical information.

In discussions with the potential users the type of delivery service was discussed. In general, the use of direct wire distribution of experimental data is not a requirement at this time. It is possible that such a service might be desirable for data from an operational satellite such as EROS, but not for AAPA. The use of airplane, U. S. Mail or hand carried delivery appears to be adequate. For at least the initial phases of AAPA a preference was expressed for hand carried or the U. S. Mails. As the organization became more familiar with the data base associated to a given industrial sector it appears that the U. S. Mail will be adequade.

4.4 User Interface

A summary of the need for the user interface primarily consists of close coordination and communication between the representative of the AAPA and the user. The AAPA representative must know the specific interests of a user and the user must be familiar with the experimental data base which is available to him. The potential user interface established during this initial study suggests that a close personal contact from a centralized system would provide sufficient feedback information about the experimental data to maintain an efficient system for both the AAPA and the user. Other types of data which could be distributed through decentralized methods can either become lost in the system or not be efficiently used by the recipient organization.

The cost for a productive unit derived from the experimental data may be considerably higher in a decentralized system, and particularly during the initial contact with the user by the AAPA, specific areas of interest will be defined eliminating the need for "shotgunning" information to a potential user. The consistent user and a growth pattern of new users will only be accomplished by providing experimental data which can be converted into some form of profit. Therefore, to swamp a user with experimental data or information will not only prevent him from obtaining this profit but will increase the cost of distribution.

4.5 Duplicate Data Bank

To a certain extent the duplication of data banks is inescapable. Of course it is in the best interest of the AAPA to minimize duplicate data banks. Even with an extensive data compaction capability, the amount of experimental data available and useful to the potential users is great. One form of decentralization does not necessarily mean duplicate data banks. This form might be regional distribution centers established on the basis of topical information or specific disciplines assigned to the regional distribution center. This configuration has the disadvantage of increased facility requirements, record keeping and a large increase in the number of secondary sources supplying experimental and supplementary data to the user. This form of decentralization tends to support and augment a particular group in a specific discipline because of their interest in this discipline and also tends to impede intraexperiment usage of experimental data.

In the present configuration at least one complete duplication will be required of the primary experimental data. The primary data will be generated by the responsible center and immediately forwarded to the experimenter. At the same time a duplicate will be made for the AAPA. The advantage of a centralized AAPA would not only be minimization of duplication but this type of AAPA could record the amounts of usage of specific data and provide direct control of it.

One of the AAPA objectives should be to see that data is only duplicated when required. For example, specific space oriented experimental data might be copied from the experimenters records for the National Space Science Data Center (NSSDC) at Goddard. In that event the AAPA should not have this data in the local data bank but should maintain adequate records that this specific data is located at NSSDC. In this event NSSDC would be considered as a secondary source and either the user would be referred directly to NSSDC, or the AAPA would obtain the data from NSSDC and add it to the work package which is required by the user.

4.6 Topical Decentralization

Topical decentralization of AAP experimental data is defined as establishing an archive which distributes data associated with specific experiments where the experiments can be classified as belonging to a particular topic or discipline. During the course of this initial study it was found that it is very difficult to assign specific S.I.C. numbers to most of the experiments on the AAP. Primarily this was due to the fact that most of the experiments are designed to obtain basic scientific or fundamental engineering facts. Since these facts become the basis for all forms of technology, the prediction that a given experiment is oriented toward a given type of industry is very difficult. In some cases (i.e., photography), however, it appears that one sector of the economy (Extractive Industry) might benefit appreciably more from a type of experimental data than other sectors. Even within the extractive industries the uses to be made of the data are quite different. For example the use of the data for minerology is quite different than the use required for petrology. Further disciplines considered along this line such as agriculture, fisheries, weather, pollution, forestry, etc. tend to increase the difficulty of dividing the experimental data into a topically decentralized form without creating an excessive amount of duplicate data banks and cost.

An advantage of topical decentralization might be that disciplines of a particular nature could be assembled at the distribution center for greater understanding of the experimental data as associated to the given topic. Upon examination of the skills required to interpret the data and the number of uses to which this experimental data is likely to the put, it is doubtful that such an advantage can exist. Certainly, interest in a broad area of industrial activity can be motivated at such a decentralized point, but it is questionable as to whether the inherent competition of industry would be satisfied.

To some extent topical decentralization is an inherent part of the existing program. The distribution of the original experimental data to the experimentor as well as to the AAPA, NSSDC, etc. will constitute decentralization of primary data. Secondary data generated by these groups should be recognized by the AAPA and duplication minimized, if possible.

4.7 Structure Recommendations

During the course of this study four configurations of structure were kept in mind for application to the AAPA. These configurations have the form as shown in Figures 26 through 29.

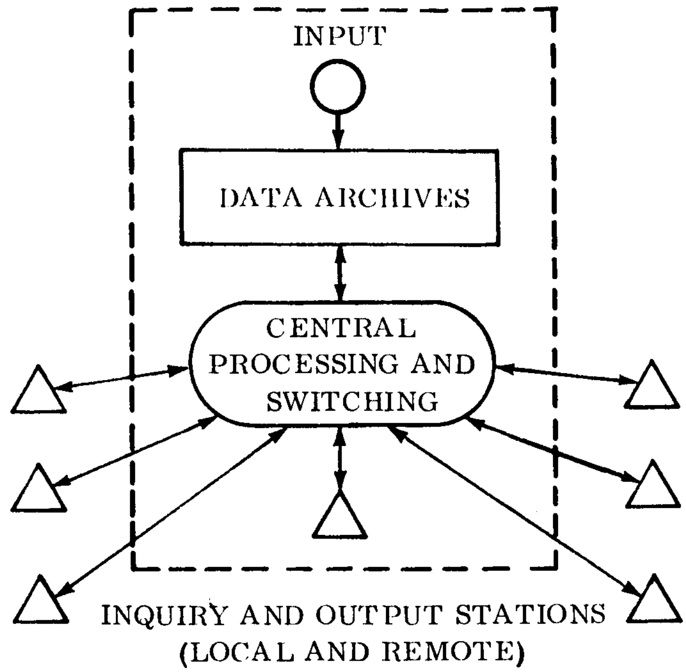


Figure 26 Centralized System with Remote Transceivers

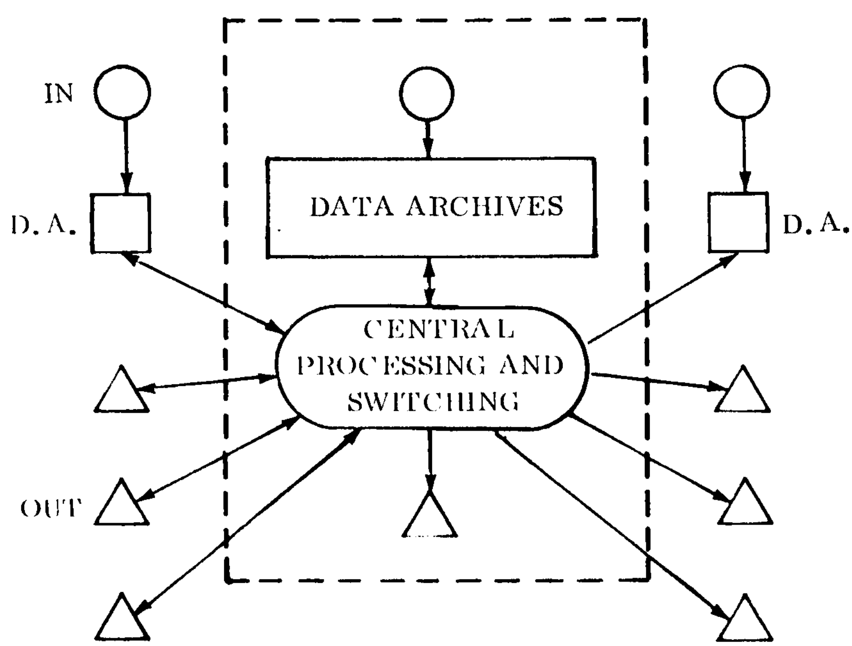


Figure 27 Centralized System with Remote Data Archives

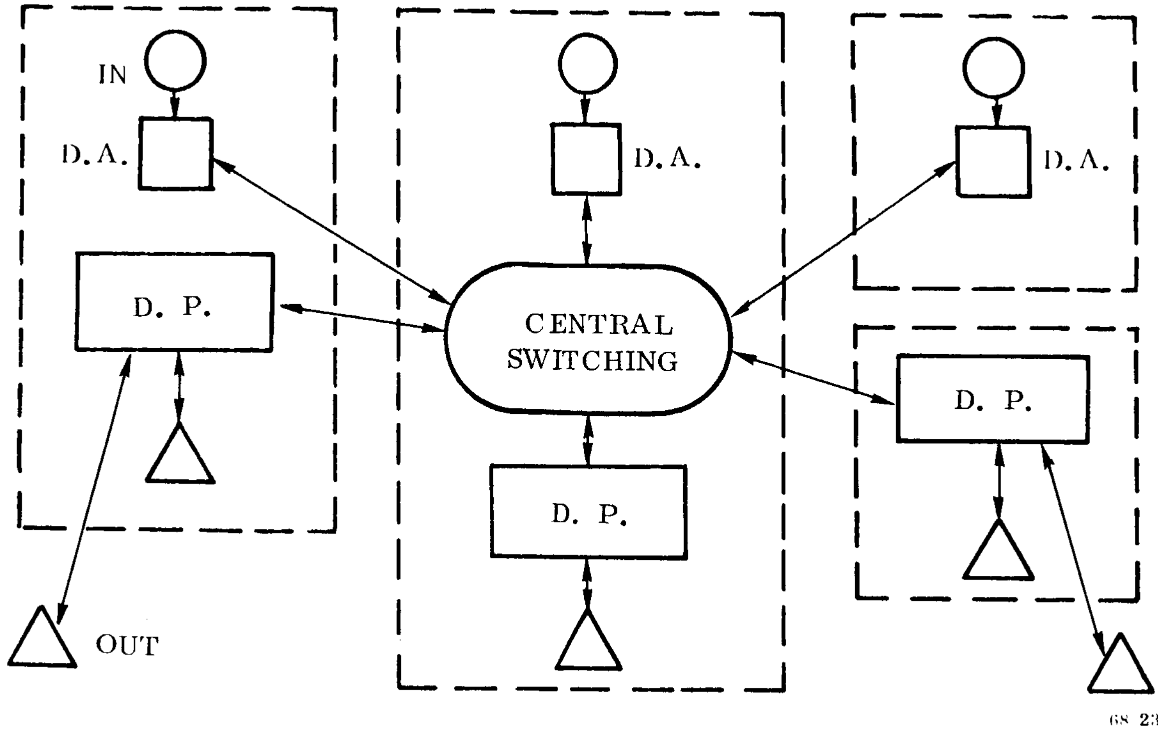


Figure 28 Decentralized System with Multiple Processors and Central Switching

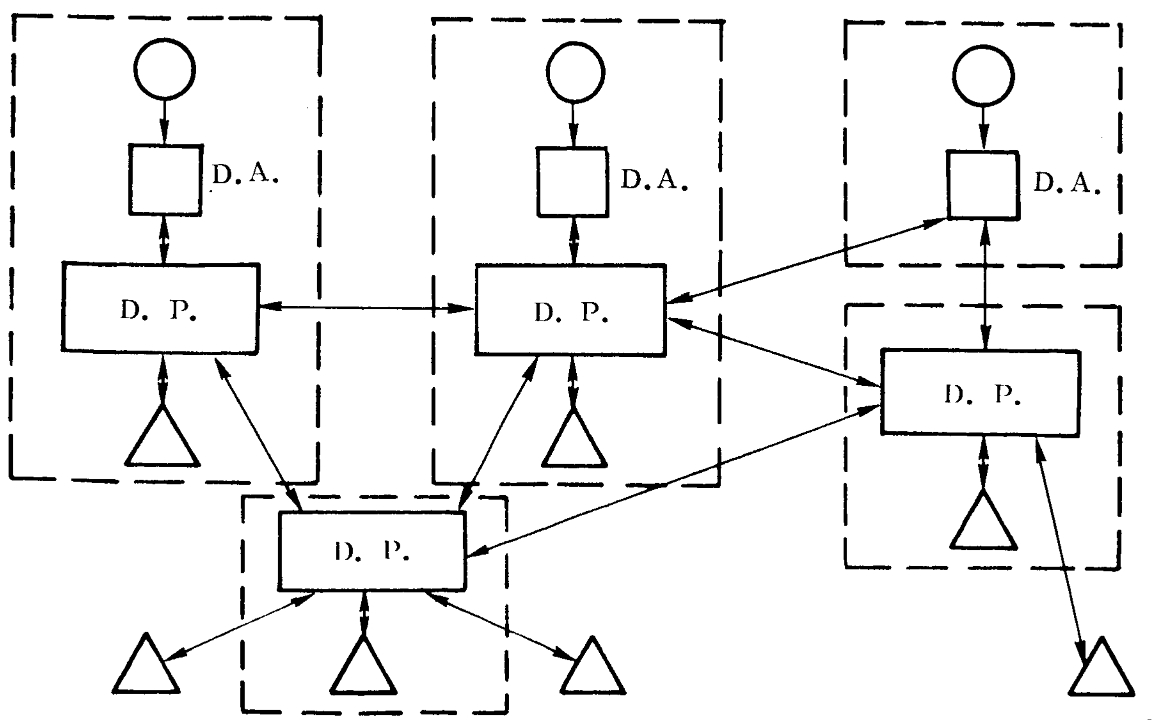


Figure 29 Completely Decentralized System with Redundant Links

These conceptions consist of several alternative structural layouts of the system depending on the degree of centralization of:

1. The data repositories and archives.
2. The processing and switching functions.
3. The input, inquiring and output stations.

After consideration of the alternatives it is recommended that a pilot system such as that described in Section 3.0 and based upon the first level flow diagram (Figure 15) be implemented. Such a system is at present a centralized system but can be expanded into a decentralized system if such action becomes warranted in the future.

5.0 FEASIBILITY OF UTILIZING THE NATIONAL SPACE SCIENCE DATA CENTER

5.1 Relationship of Scientific and Experimental Data

The data dealing with, and generated from, the Apollo Applications Program fall generally into three categories: scientific, operational and experiments related. The archival system under consideration in this study is concerned with experiments related data. The National Space Science Data Center (NSSDC) at the Goddard Space Flight Center, Greenbelt, Maryland is concerned with scientific data. Therefore, it has been necessary to investigate the mission and characteristics of the scientific data system for comparison with those of the experimental data archives system and to evaluate the advantages and disadvantages of merging the two systems, and/or to what degree.

Consideration must be given to the use differences that exist between science-oriented data and information users as opposed to engineering/industry-oriented users. It is assumed that the large quantities of experiments-related data will be used by engineering/industry, since there are a greater number of potential users in this area and their interests include the applicability of multidiscipline data or information. The fundamental differences between information use patterns of science-oriented and industry-oriented personnel have been detailed in a series of recent government-funded studies (i.e., see the volumes of DOD User-Needs Study or T.J. Allen's report on "The Differential Performance of Information Channels in the Transfer of Technology," MIT Sloan School of Management, June 1966). These studies indicate, in general, that the use patterns of industry-oriented personnel are strongly shaped by the boundaries of the "local work environment" and related factors. The boundary impedances have a tendency to make the engineer or industry-oriented personnel to turn inside his company and its relationships for data or information. The science-oriented user tends to have a different set of boundary impedances which limit his information acquisition to a specific discipline. Generally he is willing to go beyond his local work environment for data or information but tends to stay within a given discipline thereby limiting the applicability of his data or information. Therefore, it is suggested that the successful archive system for AAP which includes a data and information system must employ a technique or techniques which can reach the engineering/industry persons inside their circle.

Indications are that NSSDC is a well organized, discipline oriented, service dedicated to servicing the space science community. NSSDC has a dual importance to the AAP Data Archives Study, as follows: it provides an insight into a similar type of S/R system presently in operation within the same organizational and operational framework; and it provides information affecting the decision whether the AAP archieves should be an independent entity, or an adjunct of the NSSDC. It would be possible to model an independent system after NSSDC, innovate an entirely new archival S/R concept, and cull most favorable features from each of these. In all, five trips have been taken to the

NSSDC and the Goddard SFC to ascertain literature and examine the total data flow from the tracking station to the user.

Extensive work and time has been spent in this area because of the natural background it affords in archival systems. The equipment complement and existing facilities are quite extensive and can be further evaluated to not only use parts of this system as a model but give further consideration to the joining together of data bases in order to prevent duplicity of data storage and facility.

Those facilities contacted were geared to handle basic space science data for a select group of people. Their main objective is the discovery of scientific fact as opposed to the engineering application of experimental data. It appears that Goddard could be a user of part of this experimental data to support their scientific pursuits, but the applications aspect of the archive under consideration is in conflict with the NSSDC mission as presently defined. NSSDC reviewed the executive summary of AAP experiments, and indicated that certain experiments listed in the summary would not be of interest to NSSDC. At the present time, it is not within their charter to collect earth resource, biomedical, communications, basic biology, space vehicle technology and engineering, or navigation and traffic data. They are interested in Space Science Data, only.

From the series of visits to the Goddard Space Flight Center and the National Space Science Data Center (NSSDC), a flow diagram (Figure 10) was prepared which illustrates both how data is processed from the time it is recorded at a tracking station to the time it is stored or distributed to the experimenter, and also how information is collected from the experimenter and disseminated to interested users. It is to be noted that a distinction has been drawn between data and information. By referring to data, what is implied is that raw data which is recorded during an experiment and furnished to an experimenter. The reference to information, on the other hand, refers to reports and/or data which has been polished by the experimenter to best illustrate the results of his work. Thus, by this definition, the NSSDC is strictly an information storage and retreival system and a distribution facility to users of all classes of experimental data generated on the AAP.

The review of the AAP executive summary resulted in a requirement for prospective data that would be obtained on the AAP. This request was in concert with the Space Science communities desire for experimental data on space oriented phenomenon and did not consider experimental data of an engineering nature, which would be of interest to industry as a group, or their interdiscipline interests.

Of the experiments listed in the Executive Summary, those desired for archiving by NSSDC have previously been noted in Table IX. This included both those experiments which they definitely wanted, plus some others which were not defined, in detail, to the degree that a decision could be made.

5.2 NSSDC Operation

The National Space Science Data Center (NSSDC) was organized to support investigations in the space sciences by making available scientific data as well as facilities for working with such data. The facilities are located in Building 26 of the Goddard Space Flight Center and include space for conference and work rooms as well as space for data processing, storage, photo and data reproduction, etc. A figure showing the NSSDC is provided in Figure 30, and the discussion of their operation follows below.

As the name implies, the NSSDC deals exclusively with the collection, storage and dissemination of experimental data which is oriented toward the space sciences. In the area of space science data, there are two classes of data which NSSDC will store: viz., reduced data and final analyzed data. Reduced data is defined to mean that data which is prepared from the raw data records (furnished to the experimenter by Goddard, Houston, etc.) by the introduction of calibration factors and correlative data and by the elimination of noise and redundant information. Records prepared from the analysis of returned samples and photographs and visual data such as the photographs themselves may be considered as reduced data. The bulk of this data will contain the values of quantities which are in a usable form as a function of time, position, or some other appropriate parameter. It is from the reduced data, or the tabulations and graphs prepared directly from them, that the experimenter will develop his analysis and conclusions. In most cases the experimenter is responsible for the completion of the data reduction through contractual obligation with NASA. This contractual obligation was only recently instigated; thus in some instances, from experiments which were flown in the past, the NSSDC is having to perform some data reduction in order to get a listing of the type of data they desire.

The final analyzed data, on the other hand, is that data which the experimenter uses in his prime analysis and supports his findings. It has therefore undergone further reduction and analysis than the reduced data and includes any correlative data used by the experimenter in arriving at his results. The final analyzed data may be analog or digital like the reduced data, including such items as ionograms, magnetograms, and photographs.

When any new experiments are approved by NASA for flight, the investigator is informed via letter by OSSA (Office of Space Science and Applications) of his responsibilities for reduction, prime analysis, and delivery of certain data records to the NSSDC. The primary specification concerning the NSSDC is that roughly six months following his receipt of the raw data, the investigator must file a data analysis plan with NSSDC which specifies in what stages he will reduce his data and what he expects to furnish to NSSDC. From this data analysis plan, agreement is reached as to exactly what data the experimenter will furnish and when he will furnish it. As mentioned above, NSSDC is generally interested in only the reduced data and the final analyzed data, and usually allows the experimenter two to three years (following his receipt of the raw data) to furnish this data. NSSDC will also work through contacts at the NASA field offices before directly approaching an experimenter for data.

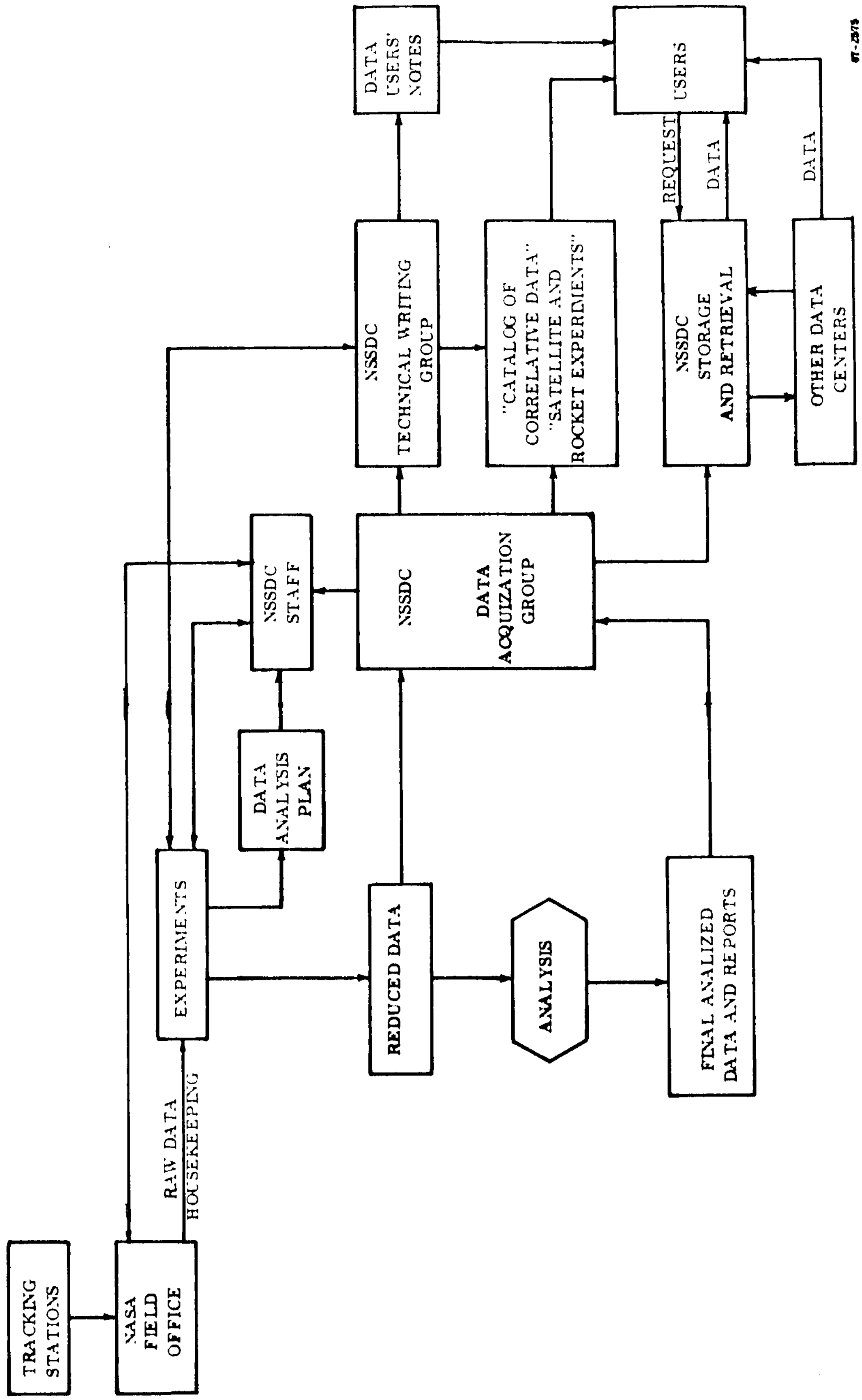


Figure 30 NSSDC Flow Plan

In addition to his data, the experimenter is also required to furnish a copy of his final analyses and reports. NSSDC maintains these reports on file, primarily for their own use in compiling the "Data Users' Notes," and in general does not reproduce them for distribution to users. Should a user desire to see these reports, they are available at NSSDC. Usually, however, all material has been published elsewhere and is available in the technical journals. To inform the users of what data is available, the NSSDC issues three publications; viz., the data catalog "Satellite and Rocket Experiments," the data catalog "Catalog of Correlative Data," and the "Data Users' Notes." The first two are published on a biannual basis and merely list available satellite data and related ground correlative data. If the correlative data is not stored at NSSDC, the catalog indicates where such data may be obtained. The "Data Users' Notes," on the other hand, are specifically designed to help potential data users decide if they can make use of the available space data. The notes are prepared by technical writers on the NSSDC staff (contracted workers) and must be approved by the experimenter before release. Included in the notes are the following: a description of the experiment and the measurements taken; the background of the experiment and the experimenter; the data reduction techniques employed; the format of the available data; a bibliography; and a list of references. The printing of the two catalogs is handled by the Government Printing Office while the "Data Users' Notes" are published at Goddard.

A major problem in the NSSDC operation is that of being able to forecast the data processing and storage requirements well enough in advance to insure adequate facilities. To date, two techniques have been employed, one of which provides a rough long-term estimate and the other of which provides a somewhat more accurate, but shorter-term estimate. The long-term estimates are based upon externally performed studies as the one recently compiled by Bellcomm, Inc. The short-term estimate, on the other hand, assesses the amount of raw data supplied to the experimenter by Goddard, Marshall, or Houston MSC and applies a general extrapolation formula. This later procedure takes advantage of the two to three year delay period between the time an experimenter receives his raw data and the time he must furnish his results to NSSDC.

5.3 Potential Function in the AAP Archives

As mentioned above, the NSSDC was organized to support investigations in the space sciences by making available experimental data in these technological disciplines. To best serve this aim, NSSDC trys to work closely with its users in an attempt to understand their needs. The NSSDC services the scientific community predominantly and thus attempts to orient its structure and "thinking" along scientific veins. The data associated with the Apollo Applications Program which would be stored in an archive, on the other hand, is more oriented toward applied research by the industrial and the educational communities.

Thus for these reasons, namely that the information is applied to different disciplines and that the users will probably require different types of information and have differing needs, it is recommended that the NSSDC and the AAP Data Archives should be separate organizations. It is further felt that closer coordination with NASA's Technology Utilization Group, the National Oceanographic Data Center, The Nation Weather Record Center, and the National Aeronomy and Space Data Center would prove of greater value than those with NSSDC. This coordination can be established within the framework of the AAP Data Archives or by a separate NASA group which also interfaces with the NSSDC and other NASA information service groups.

5.4 Estimate of Required Changes

An equipment complement for NSSDC was obtained and was reported in the Letter Progress Report No. 1, and a list of projected NSSDC equipment for the next several years is also available. However, since the charter for NSSDC specifically orients this group towards space sciences it presently appears that it is not feasible to include the AAP Data Archives within or join it directly to the NSSDC; therefore, no effort has been expended to date in assessing the equipment requirements for such a combined structure.

6.0 ARCHIVE STORAGE AND RETRIEVAL REQUIREMENTS

6.1 General

The fundamental concept of the AAPA - that of an archive storage and retrieval system - embodies a dual outlook. As an archive it must have the characteristics of a permanent repository. This normally implies measures taken to ensure the performance of source data records, including the duplication and periodic exercising of data, with strict quality and environmental control throughout. As a data storage and retrieval system, it must fulfill the single demanding requirement of being "user-oriented." This simply implies that the system provide the right data to the right person at the right time, all with a minimum of inconvenience to that person. An unfortunate, but important, aspect of this type of system in general is the common reluctance of people to use them, however, well devised. As stated by "MOOERS'LAW":

> "...an information system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it."*

In other words, the information (data) cannot merely be available; it must also be readily retrievable. This is not a characteristic of most archive systems.

Another unfortunate characteristic of most retrieval systems - and this one is no exception - is that little is known concerning the needs of users with respect to the desired data content, characteristics, and form, although this is one of the most important considerations in formulating the retrieval system. Designing a system to satisfy unstated needs may sound like an impossible task: it is not. In fact, it is the normal situation. It is accomplished by establishing intentional or implicit limitations in scope and purpose. The system must not attempt to be all things to all people. Based on a certain degree of knowledge of what kind of data will be available, and which of these data should interest the intended field of users, the system can be designed from past interests and trends shown by similar systems. As experience with the new system accumulates and the actual requirements become more precisely known, the system can alter course or expand to respond to its needs.

Conversely, the users will also adapt in time to an existing system. If the system has failed to provide its users with adequate responses, and seems to them unlikely to do so in the future, unusual demands on the system may diminish. (Mooers' Law again). Accordingly, the stability of a pattern of usage does not in itself indicate that a system is being successful in anticipating needs. Success must be related to the satisfaction and creativeness of the intended users, nor merely the current ones.

---

*Mooers, C.N., "Mooers' Law, or why some retrieval systems are used and others are not." Zator Co., Technical Bulletin No. 136 (December 1959).

Concerning the operation of the storage and retrieval system, two different strategies can be followed. One is to analyze and organize the collection with great precision in the anticipation of user questions. This is usually done by means of indexing. When a question arises, one would presumably have the pertinent records of their index entries already segregated from the rest of the collection, making retrieval rapid and routine. The second strategy is to avoid any unnecessary prior processing of records. When a specific question is received, a record-by-record search of the collection is made. The first strategy makes sense where needs can be anticipated precisely, where rapid retrieval is essential, or where processing costs (which are proportional to the number of records processed and the detail of processing) are low compared to the alternative searching costs (which are proportional to the number of records examined in a search and the frequency of searching). The second strategy makes sense where the opposite conditions exist. In practice, virtually all systems employ a blend of these strategies in a manner which hopefully achieves the best compromise between the two extremes. There are as yet, however, no quantitative guidelines for selecting the best means of service activities for a system, or for determining the precise emphasis to be given each aspect of record processing.* Nevertheless, even on the basis of some of the more obvious AAPA characteristics - very large number of records, relatively low frequency of use, no real-time demands, reasonable tolerance for trail-and-error retrieval - it can be seen that the AAPA develops a much stronger rationale for employing the second strategy than the first.

In order to better analyze the overall requirements for this archive storage and retrieval complex, it is convenient to consider separately the five basic functions of such a system. They are:

1) Acquisition - including the decision of which types and forms of data to acquire.

2) Surrogation - including the analysis, cataloging, indexing, and abstracting of data

3) Announcement of the availability of acquired data items to the field of intended users.

4) Data Management - including data file conversion for storage, storage, retrieval, and the processing and reproduction required for dissemination.

5) Dissemination - physically or electrically transferring data to the user.

The acquisition function was discussed in Section 3.0 The characteristics of the anticipated experiments data were reviewed and estimates of data volume and rates of accrual made. Using these results and the tentative experiments data flow plan, the

*"Information Storage & Retrieval," Lipety, Ben-Ami, Scientific American, September 1966.

input boundaries and acquisition responsibilities of the AAPA were defined. Beyond this is the important, though mundane, 'sine qua non' operation of implementing a working acquisition scheme. Unless successful relationships are established with the various data purveyors - the PI, GSFC, MSC, MSFC - the acquisition paths which are so readily flow diagrammed may never become completely operational.

Surrogation is defined as the process of substituting for an item something that will represent it or stand in its place for various purposes. It may include one or more of the processes of cataloging, abstracting, and indexing. In an information or data retrieval system, a surrogate may be an accession number, a title, a citation, an extract, or an abstract. The concepts of primary, secondary, tertiary, and quarternary data defined and described in Section 4.6 establish the breadth of surrogation required of the AAPA. As it is described in the first level system flow of Section 4.0, this normal function of surrogation has been distended to encompass the announcement function as well. It can be readily seen that the surrogation function is primarily an intellectual operation which directly entails little or no hardware. It does, however, entail much of the software aspect of information storage and retrieval, particularly concerning the subject of indexing for subsequent retrieval. Some presently implemented software systems pertaining to indexing and retrieval are presented in Section 6.2.

The data management function essentially pertains to all data handling, accomplished internally within the AAPA. It can be more effectively discussed in terms of its main constituent parts of storage and retrieval hardware and techniques, and storage and retrieval software. As mentioned above, the software aspect is discussed in Section 6.2. A state-of-the-art review of storage and retrieval hardware and techniques is given in Section 6.3.

6.2 Software Considerations

Several indexing and retrieval programs have been written for general-purpose computers. The following types are covered:

1) Permuted Title Index Programs - Significant words in document titles are arranged alphabetically and published in an index.

2) Search Programs - Parameters of a user-initiated question are compared with a stored file in order to obtain information relevant to the question.

3) Selective Dissemination Programs - Index information associated with incoming documents is compared with profiles of users interests, and a reference is sent to the user where a close match occurs.

6.2.1 Permuted Title Index Programs. - Key Word in Context (KWIC) Index Program produces a permuted title index, and, if the user requests, a bibliography and author index. In addition, a bibliography tape is produced which may be used for preparing other indexes. Other options include a choice of three page formats for each of these printouts, and a provision for frequency counts on all words appearing in the index or excluded from the index. IBM, for example, has a variation of this program for almost every computer that it manufactures.

6.2.2 Search Programs. - Here a distinction is made between card-oriented search and report generation programs and generalized information storage and retrieval search programs. Differences are essentially one of degree.

- EDP Search and Report Programs. - These programs tend to be less flexible, concentrating on scanning of fixed data fields for specific purposes. Searching is only one part of a much larger function, including file maintenance. This integrated aspect makes these programs less flexible and harder to adapt.

- Information Storage and Retrieval Search. - IS&R search programs are a much more flexible kind of file updating and fact retrieval program, and the search strategy tends to be more complex. Also, the IS&R program stresses the man-machine communication aspects of information retrieval. EDP programs are set up to ingest and retrieve data and produce reports on a regular basis for jobs not requiring restructuring. IS&R programs will accept a tremendous variety of inputs and requests with no reprogramming and will deliver new types of answers in a very short period of time. Finally, IS&R searches may be performed to satisfy data requirements which are periodic or varied, rather than continuous and standard.

- RCA RECOL (REtrieval COmmand Language). - This is a fact retrieval language which is intermediate. The language is a general interrogation scheme for small to medium sized files requiring a minimum of programming effort. The data format of the file is specified by the insertion of a table of record contents within the control program at the time of assembly.

  Up to five questions may be asked in a single pass through the file. There are five search orders which may be combined in various ways to form one question. The many combinations of five search orders insure considerable flexibility in output format.

- IBM IPS (Information Processing System. - IPS is a large and flexible system that permits a wide variety of queries. The two basic components are a File Maintenance Program and in Information Retrieval Program. Retrieval queries are stated in the language of IBM IPS which permits many variations in the queries.

The system maintains a library of query programs as well as a table to describe the format of each file. The query, entered in the form of punch cards, initiates the proper routine to translate the query. The output is printed in a user-specified format.

A control program handles the processing functions being performed by the system.

6.2.3 Selective Dissemination Programs. - IBM SDI (Selective Dissemination of Information) establishes and maintains a current list of people who join the SDI system. Each person is represented on magnetic tape by a list of keywords originated by him, that describes his particular work interests.

As new documents are published, a list of keywords from each document is entered in the system. The computer watches the keywords of each new document against the users records, punching out a notification card only for users potentially interested in a particular document.

In summary, it is clear that there is no one complex of software, nor is there one piece of hardware, which can perform all of the major functions involved in a total information storage and retrieval system. Such a system requires unrelated pieces of equipment tied together by a set of procedures, plus computer programs, such as the ones described above.

The organization of the software needed to implement information storage and retrieval functions does not necessarily closely parallel the functions listed. It is rather more centered around conversions to common file structures, operating system conventions, standardized program and data structure interfaces imposed by the various multilevel language translators used, the characteristics of the various categories of storage and communications used in implementing the system, and intersystem interfaces introduced by the various intercommunicating user installations. In effect, systems integration will be accomplished at different levels within a given facility and between facilities, depending on the frequency of certain interactions, the historical precedence of particular hardware-software systems, the importance attached to specific modes of operation, and the relative difficulty between making changes in subsystems providing new interfaces between them.

6.3 Storage and Retrieval Devices and Techniques

The types of data to be stored were discussed in Section 3.0 of this report. With rate exceptions, all of the data which must be stored in the AAPA will be in, or can readily be converted to, either digital or image form. In most cases, the choice between these two forms for storage is obvious, although in some cases, such as telemetered versions of image data, it is not. Within the scope of this investigation, data received in digital form will be assumed to be stored in digital form, while that received as images will be stored in image form. The received image data will consist of charts, plots, tabulations, typed reports, and a sizeable volume and variety of photographs. Although the common storage media for most image data is one of the several versions of microfilm, this is not suitable for archival storage of photographs. The reason for this is that any image degradation, including resolution loss, encountered during the microfilming process is considered a loss of source data. With the high quality, high resolution, photography scheduled for AAP, such a loss would be inevitable.* The actual photographic film, therefore, must itself be retained as the archive format. The AAPA should archive second or third generation quality photographs, all originals being retained at MSC. Within the AAPA, it would then be feasible to conceive a photographic hierarchy, the lowest order being a microfilmed version. This version is the one best configured for normal storage and retrieval handling, and is also convenient for the isolated production of relatively low quality copies. When higher quality copies are required, they may of course be made available from the second or third generation masters.

The following is a review of some current storage and retrieval devices and techniques - both digital and image - which have been reviewed for their possible applicability to AAPA.

6.3.1 Chemical Storage. - The NCR photocromic microimage process PCMI provides the capability of high density document of photographic storage by utilizing both microfilm and photochromic materials. One advantage of this technique is that image reductions on photochromic reductions are erasible and therefore convertible before final recording on microfilm. Linear reductions up to 200:1 are feasible.

A microimage system would include the following elements:

- PCMI Camera/Recorder - This device will reduce microimage from 35 mm microfilm onto photochromic film.
- Off-line refrigerated file - PCMI master plates will be stored in this file until required for updating.
- Contact printer - This unit will produce micro-image cards from the microimage master film prepared by the PCMI Camera/Recorder.
- Microimage viewer - For viewing microimage cards.
- Hard copy printer - For hard copy printout of enlarged microimages.
- Microimage card file - For storage of microimage cards.

*"Information Content in Space Photography," Fordyce, S.W., October 22, 1965, estimates the information content per picture for recon. camera at $9x10^9$ bits.

Mechanization of the NCR/PCMI process requires that the original document be first transferred to high quality conventional microfilm. Properly filtered, near-ultraviolet radiation is then directed through the transparent microfilm and into the microimage optics where it is focused on a photochromic plate. This forms a miniature image on the photochromic coating such that 3200 microimages, each representing an 8-1/2 x 11 in. page reduced 150 times, can be placed on a 4 x 6 in. photochromic plate. This plate will then be used for producing duplicate microimage cards on photographic film for dissemination. Figure 31 shows the basic operation of an NCR/PMI microform processing center.

The economic advantage of the PCMI system exists where there is a requirement for a large number of microimage copies since the cost of creating the photographic master is appreciable.

6.3.2. - Magnetic Storage. - Because of its desirable characteristics, magnetic storage has been utilized extensively in information handling. Most appealing are its erasibility and freedom from processing, allowing instant use of information after recording. Such characteristics have encouraged the development of many forms of magnetic storage for various applications. The forms considered here are:

- Magnetic Core
- Magnetic Thin Film
- Magnetic Drum
- Magnetic Disc
- Magnetic Card
- Magnetic Tape

Magnetic core and thin film storage devices are two types most commonly employed in third generation computers for internal, direct address storage. Figure 32* shows a general comparison between these and other currently employed direct access storage devices, whereas Figure 33* shows the general cost characteristics of magnetic core storage devices. Additional cost information compiled by UACSC during the course of this study and presented in Table XIII shows a somewhat higher cost trend for these memory devices. Because of their relatively high cost per bit, none of these direct-access storage devices are considered to be applicable to the AAPA mass storage requirements.

---

*"Memory Systems Comparison," Weniger, K., The Electronic Engineer, May 1967.

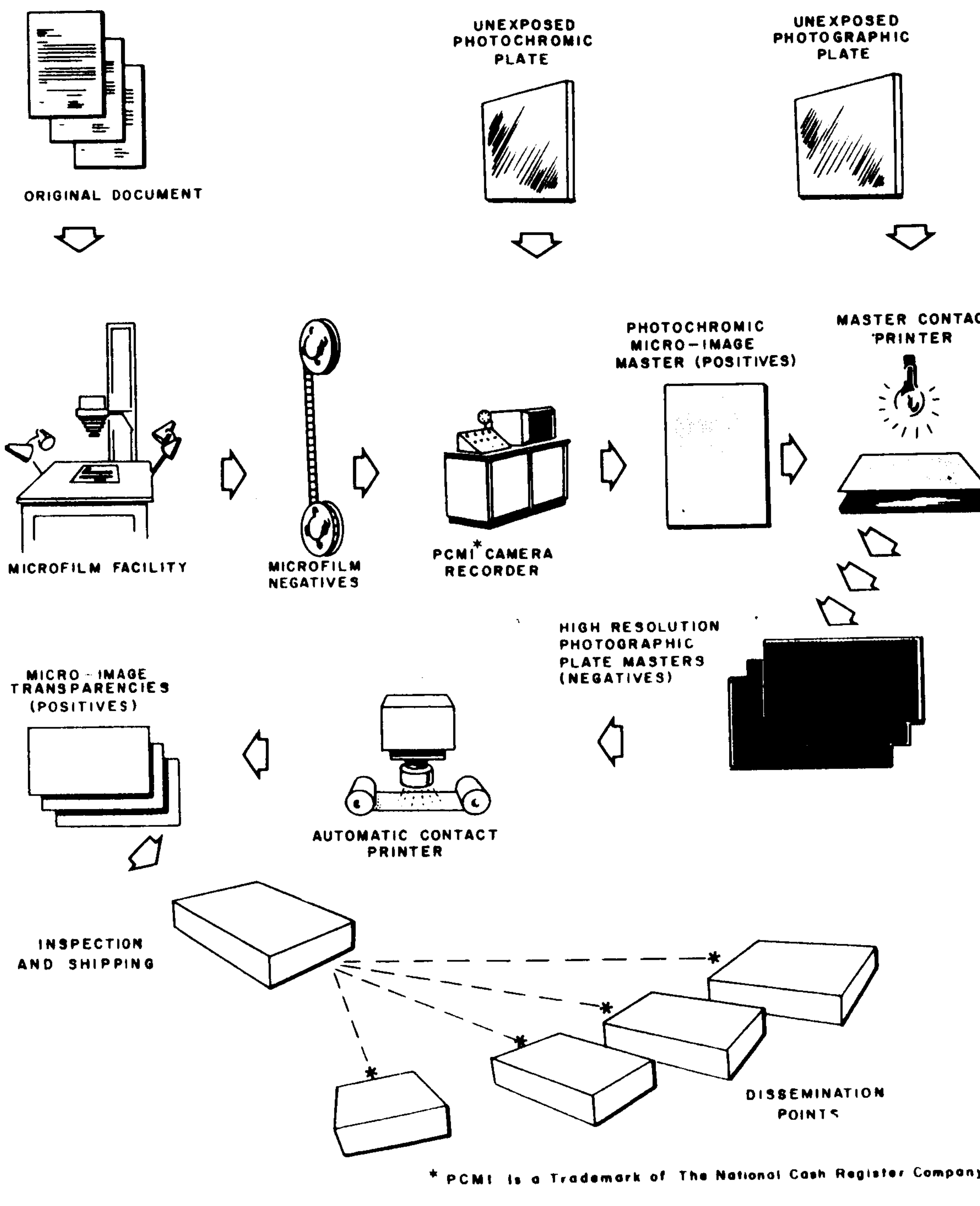


Figure 31 NCR Microform Processing Center

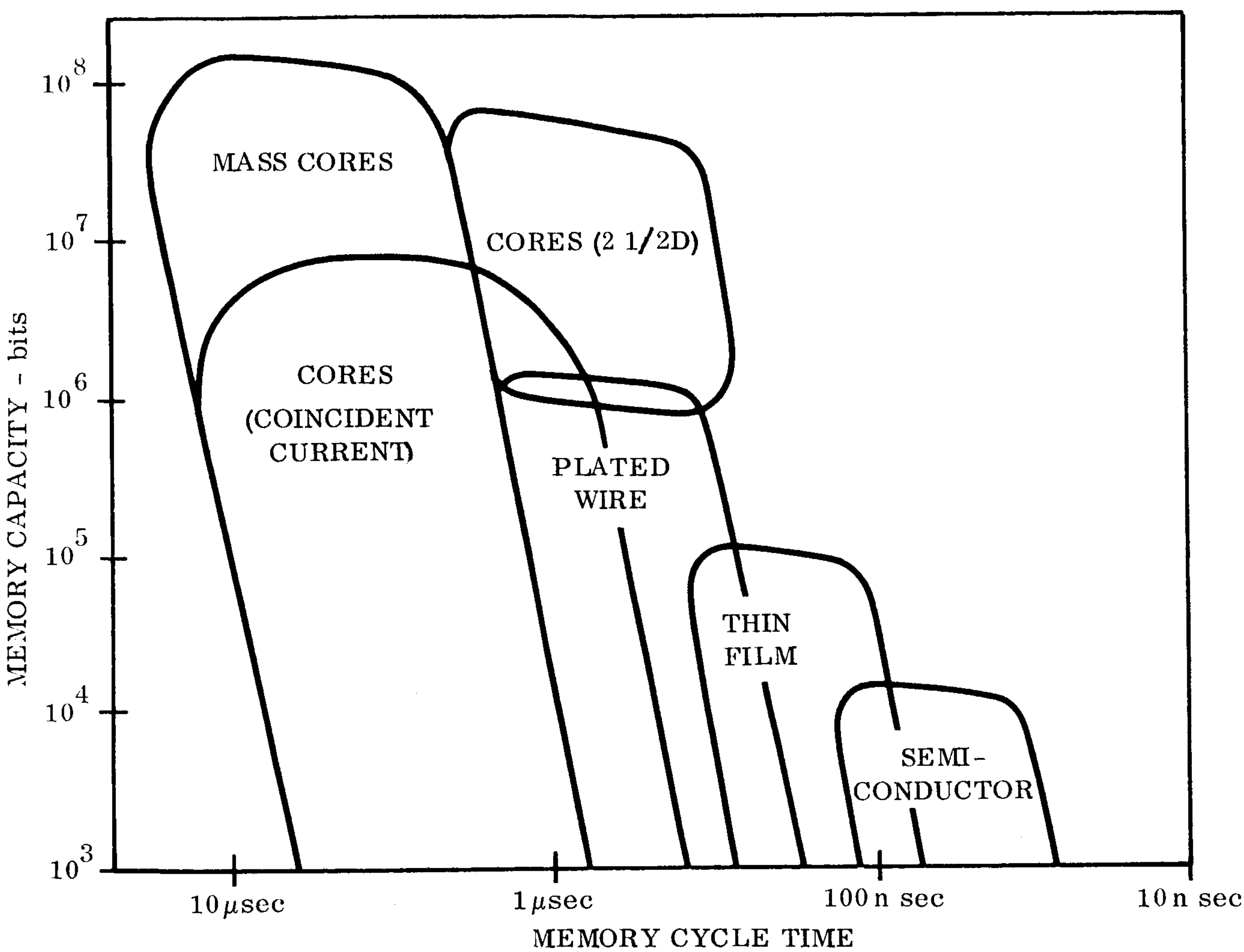


Figure 32 Comparison of Common Direct Address Storage Devices

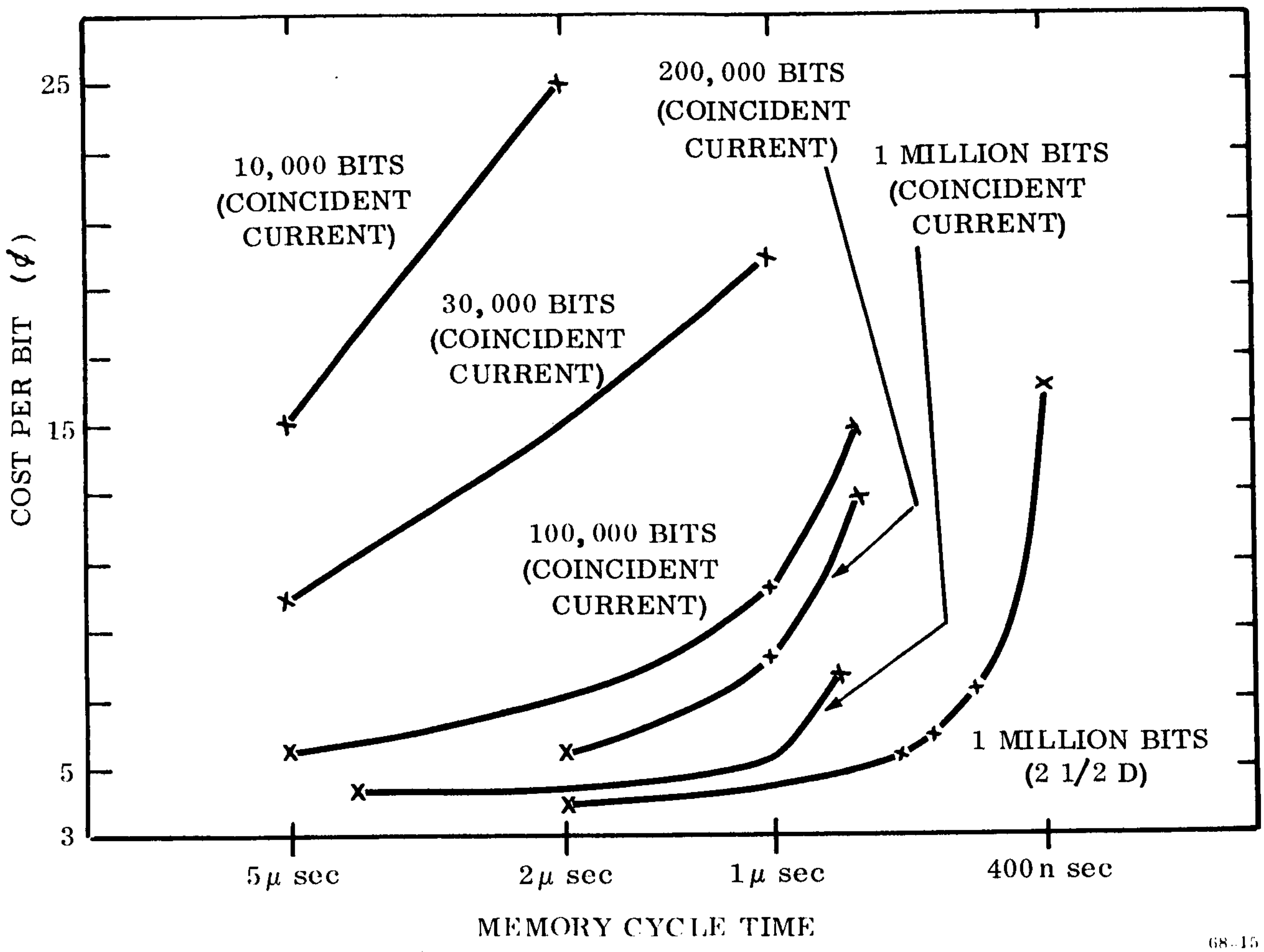


Figure 33 Magnetic Core - Cost Characteristics

TABLE XIII
MAGNETIC CORE AND THIN FILM MEMORY
COST CONSIDERATIONS

	Range Cycle Time (μs)	Average Maximum Memory Capacity (K Bits)	Average Cost of Maximum Memory Capacity	Average Cycle Time (μs)	Average Cost Per Bit (Cents)
Business Data Processors					
Core Memory	0.0-0.9	3,646	$1,281,276	.79	35.1
	1.0-1.9	1,326	365,276	1.34	27.5
	2.0-2.9	1,782	558,773	2.24	31.4
	3.0-4.9 (1)	1,816	610,603	3.93	33.6
	5.0-6.9 (2)	720	325,035	5.95	45.1
Thin Film Memory	0.0-0.9	256	103,780	.60	40.5
	1.0-1.9	128	43,940	1.20	34.3
Medium and Large Scientific Processors					
Core Memory	0.0-0.9	6,576	2,019,305	.79	30.7
	1.0-1.9	12,640	3,156,916	1.19	25.0
	2.0-2.9	5,301	1,166,300	2.00	22.0
	3.0-3.9	512	135,400	3.60	26.5
	4.0-4.9	1,950	625,000	4.80	32.1
Thin Film Memory	0.0-0.9	24,960	6,860,300	.60	27.5

The most common form of magnetic storage is magnetic tape. It is relied upon by the computer industry for bulk storage and even with recent developments, is still the most economical form of erasible storage available. Its use has been challenged in various applications because of the development of faster random access type devices, such as magnetic discs and drums. However, due to the increased cost of these, tape storage is still extensively used. A common practice is to use tape storage in combination with some form of random access device. In this manner, a compromise is reached between access speed and cost. Files of up to 100 million characters can be stored on discs and any larger files on magnetic tape. In this concept, the more active portions of the file data are recorded on the magnetic discs while the less frequently used data is placed on magnetic tape.

In the case of AAPA, the overall data base is obviously not of sufficiently high usage to warrant the employment of discs or drums as an archive storage device. These will, however, be required to implement the retrieval operation, which requires the readily accessible storage of large blocks of surrogate data, including the entire active system index, thesaurus, catalog, all or part of the data abstracts. etc. Table XIV contains storage characteristics for some representative high capacity disc and drum systems of the type which may be required for a later generation of AAPA.

6.3.2.1 Magnetic Card Storage. - Information storage and retrieval requirements cannot in most cases be defined to the extent that exact equipment characteristics can be specified. As a consequence, manufacturers usually try to include as much flexibility as possible in their equipment designs. Requirements for mass storage emphasize capacity and, in many cases, erasibility. Initially, magnetic tape was the only storage medium that could economically provide these storage characteristics. However, magnetic tape was dependent upon the roll storage concept which proved to be ineffective in handling large files unless sequential access or considerable batching was employed. Consequently, it became evident that magnetic tape storage lacked several characteristics desirable for mass storage.

Unit record storage, in contrast to roll storage, is applicable to operations which require considerable flexibility in file organization and information manipulation. In recognition of these characteristics, attempts have been made to combine the storage advantages of both magnetic tape and unit records. The results of these efforts have been the present magnetic card handling devices. One such unit is discussed below.

RCA Random Access Mass Memory System, Model 3488-2 (formerly RACE)

The Mass Memory System (MMS) accepts data from the computer one word at a time and stores the data on magnetic cards at a location specified by the Computer. When interrogated by the Computer, the data is transmitted from the mass memory to the Computer one word at a time. Each data word is stored on the card as part of a block of 875 words. Every block of 875 words on a card may be randomly accessed.

TABLE XIV

STORAGE CHARACTERISTICS OF SOME HIGH CAPACITY MAGNETIC DISC AND DRUM SYSTEMS

Entry Number	Company	Model Number	Unit Surface Diameter	Number Units Per Controller	Unit Capacity (Million Bits)	Record Density (BPI)	Track Density (TPI)	Average Access Time (MS)	Transfer Rate (1000 CPS)	Purchase Prices			Capacity Per Controller (Million Bits)	Cost Per Bit (Cents)
										Unit	Controller	Gross		
1	Librascope	L 4800	6-48" Discs	16	409.7	900	48	35	141	$200,000	$200,000	$3,400,000	6,555	.0519
2	Burroughs	9375-3	Disc	50	200.0			60	395	28,800	12,000	1,466,400(1)	10,000	.0147
3	Burroughs	9372-7	Disc	50	160.0			23	377	57,000	12,000	3,174,000(2)	8,000	.0397
4	Digital Development	7302-4	4 Discs		15.36	1000	50	8.5	300					
5	BMA	MS 412	4-12" Discs		14.3	1200	30	16.7						
6	Magne-Head	134-512	4-13" Discs		8.19	548	40	8.5						
7	Data Disc	F6	12" Disc		6.0	3333	33	16.7						
8	Bryant 4000	2600-6 Module	26" Disc	6	12,500	1500	128							
9	Bryant	4000 C-2A	39" Disc	8	3,800	800	128	122.5	458-1459	171,090	50,000	1,418,720	30,400	.0047
10	Data Products	5085	31" Disc		3,500	2000	53.4	111						
11	Honeywell	262	Disc	4	1,800			80.8	196	288,000	26,100	1,178,100	7,200	.0164
12	IBM	2302-4	Disc	4	1,794	1060	100	182		$344,830	$26,130	$1,415,450	7,176	.0197
13	Control Data	814	26" Disc	8	1,208	1070	50	91	196	241,500	25,000	1,957,000	9,664	.0203
14	IBM	2314	14" Disc	1	1,656	NA	100	87.5	312	$244,440	(3)	$ 244,440	1,656	.0148
15	Control Data	854	14" Disc	8	49.2	1105	100	97.5	208	23,000	$25,000	209,000	394	.0539
16	General Electric	DSU 160	16" Disc	8	47.16	3200	80	91	156	25,510	24,000	228,080	377	.0676
1	UNIVAC	1782	24" Drum	8	75.5	691	--	17	1440	117,200	182,515	1,120,115	604	.1858
2	BRYANT	185 1024	18.5" Drum		47.59	800	28	16.7	347					
3	Magne-Head	18-1024	18" Drum		43.4	750	50	16.7						
4	Vermont Research	232A	20" Drum		40.96	637		16.7						
5	IBM	2301	11" Drum	1	32.76	1250	80	8.6	1200	96,000	108,930	204,930	33	.6255
6	UNIVAC	FAST-RAND2	24" Drum	8	793	1000	106	126	156	164,640	51,060	1,367,180	6,344	.0216
7	Bryant	Phd 340	20" Drum	8	346	1000	128	59	1200	65,000	50,000	570,000	2,784	.0205

(1) Total cost includes cost of 1st 100 million bytes of storage (four units) with electronics unit at $129,600.
(2) Each electronics unit purchase price is $31,200. Maximum of 10 electronic units per controller. 5 unit devices per electronics unit.
(3) Controller price is included in unit price.

The data block length is arbitrary and may contain a minimum of one word or a maximum of 875 words. Access to words in a block is serial. Every card and block within the memory may be randomly accessed. The cards are contained in magazines which may be removed and replaced with other magazines. The magazine containing the cards and the mechanism for selecting cards and recording or reading them, along with the necessary logic circuitry and power supplies, are contained in a single unitized structure.

Card Characteristics

Card Size	4-1/2 x 16 in.
Channels	128
Bits/Channels	8750
Recording Density	625 bits/in.
Card Capacity	1,120,000
Card Material	Magnetic oxide coated Mylar

256 cards are normally placed in a magazine store. Each magazine can store as much data as 2-1/2 reels of magnetic tape. Additional characteristics of the MMS are given in Table XV, along with comparable units of other manufacturers. It is not believed that the AAPA, as presently conceived, warrants the employment of these type units.

6.3.2.2 Magnetic Tape Storage. -

6.3.2.2.1 Digital Storage. - The applicable references of Appendix E discuss the gamut of pros and cons related to the employment of digital magnetic tape as an archival storage media. The key fact remains that magnetic tape is the primary digital storage media employed throughout every phase of the AAP data flow. Most phases of all experiments data processing, including the blocking of data, are geared to tape. The AAPA will receive its digital data on tape, and will eventually retire it on tape. For the initial configuration, at least until its operation and growth trends become better defined, magnetic tape seems to be the logical choice for archiving digital data in the AAPA.

Although the technology of magnetic tape recording has been making steady advances - primarily in reliability - during the past decade, one recent development offers a significant improvement to the tape medium itself. A new, high-performance magnetic recording tape incorporating chromium dioxide as the recording medium has been announced by the DuPont Company. Trademarked CROLYN, this tape utilizes DuPont's proprietary chromium dioxide, incorporated in specially formulated binders, coated on Mylar$_{TM}$ polyester film. Chromium dioxide has a higher magnetic moment per unit volume than gamma iron oxide typically used in conventional magnetic tape.

TABLE XV

STORAGE CHARACTERISTICS OF SOME HIGH CAPACITY MAGNETIC CARD/STRIP SYSTEMS

Entry No	Company	Model No.	Recording Medium & Size (Inches)	Number Units Per Controller	Unit Capacity (Million Bits)	Cartridges Per Unit	Cartridge Capacity (Million Bits)	Cards Per Cartridge	Tracks Per Card	Track Length (Bits)	Record Density (BPI)	Ave. Time to Access (MS) New Track	New Strip	New Cartridge	Transfer Rate (1000 CPS)	Purchase Prices Unit	Controller	Gross	Capacity Per Controller (Million Bits)	Cost Per Bit (Cents)
1	RCA	3456-2	CARD-16 x 4 1/2	6	4,774.0	16	298.3	256	128	9,100	700	20	460	460	60	$127,500	$22,500	$1,052,500	36,192	.0028
2	RCA	[illegible]-11	CARD-16 x 4 1/2	8	4,485.0	8	561.6	256	128	17,112	1,400	20	355	385	70	130,000	25,500	1,065,500	35,904	.0030
3	IBM	2321 DATA CELL	STRIP-12 7/8 x 2	8	3,200.0	10	320.0	200	100	16,000	1,780	95	175	387	55	132,400	33,220	1,092,420	25,600	.0042
4	RCA	3456-1	CARD-16 x 4 1/2	8	2,387.0	8	298.3	256	128	9,100	700	20	385	385	60	65,000	32,500	552,500	19,096	.0029
5	NCR	353-5	CARD	1	497.7	1	497.7	384	144	9,000	936	24	90	NA(2)	50	63,000	(1)	63,000	498	.0127

This characteristic results in 1) higher signal output at the same degree of resolution and 2) better resolution at a given level of signal output. In particular, CROLYN has shown the same performance at 1600 bpi that iron oxide shows at 800 bpi.

CROLYN can be used interchangeably with iron oxide tapes in many applications, but greater performance benefits are usually obtained on recording equipment adapted for use with the new tapes. DuPont reports that several manufacturers - particularly in the computer and the industrial and home video areas - have been actively working either to modify existing equipment or to develop new equipment to take advantage of the superior properties of CROLYN.

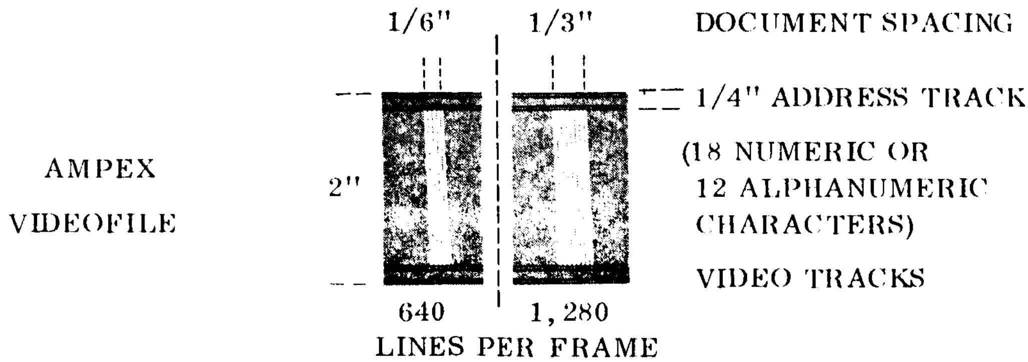
6.3.2.2.2 Image Storage. - The AMPEX VIDEOFILE system stores conventional file documents as television recordings on two inch magnetic video tape. Each paper, photo, drawing, graph or document is recorded through a camera or scanner onto the video tape. This can be done alphabetically, numerically, or chronologically, exactly as with paper files. Each segment of tape, housing one document (in an alphabetical or numerical file) or a series of related documents, can be considered as a magnetic file folder. An operator puts the identification address code on the documents or directly into the system.

To retrieve a file, the requestor dials its address number on a telephone dial or desk-top keyboards. The television recorder, which may be located anywhere in the building, receives the electronic command and searches the tape to locate the right file. When the file is pinpointed it is copied into the buffer, the original remaining in the master file for others to use. The recorder is then free to answer other requests. Average time of the whole operation is under one minute. The requestor can study all the documents in the file on his desk screen, and if he chooses, make a printed copy of any page.

The arrangement of the video and address tracks recorded on Videofile tape is shown in Figure 34.

Storage resolution is fixed by the number of lines per frame. The 1280 lines per frame, which is for use with 8-1/2 x 14 in. copy, provides a line width of 14 mils on the requestor station TV monitor, giving a resolution of slightly greater than 70 lines per in. (Resolution is also limited at about that level by the monitor itself.)

Videofile is basically a modular system which can be constructed in various arrangements of the following building blocks: filing stations, intermediate storage stations, sorting stations, master file station, and retrieval (printer or monitor) stations. The main advantages of Videofile over other image storing devices, primarily microfilm, are:



FRAME SIZE, RESOLUTION & CAPACITY RELATIONSHIPS

FRAME SIZES	MAXIMUM DOCUMENT SIZE PER FRAME*	LINEAR TAPE LENGTH PER FRAME	DOCUMENTS PER REEL		
			14" DIAM	10½" DIAM	8" DIAM
640/30 Frame	5½ x 8½"	1/6 Inch	500,000	250,000	125,000
1280/15 Frame	8½ x 14"	1/3 Inch	250,000	125,000	62,500

*To resolve 10-point (or larger) type

68-19

Figure 34 Videofile Image Storage Arrangement

1) File flexibility - ordered, random, or a combination of the two can be employed. Files can be readily updated, modified, or reordered.

2) Continuous accessibility of documents - the use of a video disc buffer file to service individual requests means that the master record is always kept in file for accessibility to other users.

3) Multiple and remote viewing.

A comparison of Videofile with some common microfilm techniques is given in Table XVI. A key difference between the two basic techniques, however, is afforded by the publication Ampex Readout: "Microfilm is also visual storage, but it is primarily intended for archival storage, whereas Videofiling is active day-to-day filing and retrieval."

The first Videofile system was delivered in 1966 to NASA at Huntsville, completing a $1,080,000 contract, where it is presently being adapted for service with PRINCE/APIC. One study completed at this location has been an evaluation of storage and retrieval cost versus number of documents stored per day. The results of this study are shown in Table XVII.

TABLE XVI

VIDEOFILE VS SOME COMMON MICROFILM TECHNIQUES

CHARACTERISTICS	MICROFICHE	APERTURE CARDS	RECORDAK MIRACODE	AMPEX VIDEOFILE
NORMAL STORAGE CONFIGURATION	4" x 6 " SHEET FILM IN PAPER JACKET	35-MM. FILM FRAME IN 3-1/4" x 7-3/8" CARD	100-FOOT REEL 16-MM. FILM	7,200-FOOT REEL 2" WIDE
SIZE OF STORAGE CONTAINER	30 JACKETED SHEETS PER INCH OF FILE	125 CARDS PER INCH OF FILE	4" x 4" x 1"	16" x 15" x 3"
STORAGE CAPACITY, 8-1/2" x 11" DOCUMENTS (REDUCTION RATIO)	60 PER SHEET (20:1)	UP TO 8 PER CARD (24:1)	2,000 PER REEL (24:1)	250,000 PER REEL (1,280 LINES PER FRAME)
MAXIMUM DOCUMENTS PER CUBIC INCH	75	42	125	350
SEQUENTIAL SEARCH RATE	(MANUAL)	UP TO 33 CARDS PER SECOND	200 DOCUMENTS PER SECOND	1,050 DOCUMENTS PER SECOND

TABLE XVII

SAMPLE VIDEOFILE STORAGE AND RETRIEVAL COST

DOCUMENTS PER DAY 20PP/DAY	STORAGE MEDIUM	TYPE OF COST	COST PER DOCUMENT AFTER YEAR(S) 1	2	3	4	5
STORAGE COST							
3200 PAGES 160	HARD COPY	FIXED AND VARIABLE	$ .027	$ .025	$ .024	$ .0235	$ .023
	VIDEOFILE (2 SHIFTS)	FIXED AND VARIABLE	16.00	9.00	6.50	5.00	4.50
		VARIABLE ONLY	1.60	1.60	1.60	1.60	1.60
80	HARD COPY	FIXED AND VARIABLE	.027	.025	.024	.0235	.023
	VIDEOFILE (1 SHIFT)	FIXED AND VARIABLE	32.00	17.40	12.40	9.90	8.40
		VARIABLE ONLY	2.40	2.40	2.40	2.40	2.40
RETRIEVAL COST							
50	HARD COPY	VARIABLE	$ 1.87	$ 1.87	$ 1.87	$ 1.87	$ 1.87
	VIDEOFILE (1 SHIFT)	FIXED AND VARIABLE	51.00	27.50	19.50	15.50	13.15
		VARIABLE ONLY	3.12	3.12	3.12	3.12	3.12
100	HARD COPY	VARIABLE	$ 1.42	1.42	1.42	1.42	1.42
	VIDEOFILE (1 SHIFT)	FIXED AND VARIABLE	25.80	13.80	9.85	7.84	6.65
		VARIABLE ONLY	1.84	1.84	1.84	1.84	1.84
200	HARD COPY	VARIABLE	$ 1.37	1.37	1.37	1.37	1.37
	VIDEOFILE (1 SHIFT)	FIXED AND VARIABLE	13.50	7.00	5.40	4.10	3.50
		VARIABLE ONLY	1.16	1.16	1.16	1.16	1.16
525	HARD COPY	VARIABLE	1.30	1.30	1.30	1.30	1.30
	VIDEOFILE (1 SHIFT)	FIXED AND VARIABLE	5.00	2.86	2.10	1.70	1.48
		VARIABLE ONLY	.57	.57	.57	.57	.57

6.3.3 Photographic Storage. -

6.3.3.1 Digital Storage. - Two recent high capacity (10^{12} bits) devices are discussed for digital storage.

6.3.3.1.1 IBM Photo-Digital Mass Storage Device. - This mass storage system, which can contain up to seven modules of one-third of a trillion bits each, uses a new combination of technologies, including: 1) electron beam recording of digital data on the photographic film chips, housing chips in small plastic cartridges, 2) pneumatic transport of cartridges to recording and reading stations, and 3) optical reading of data at a rate of 2.5 million bits per second.

The data recording mechanism of the IBM-developed system includes an electron beam recorder and film chip developer. It receives unexposed chips, records data on the chips and automatically develops them. The finished chip then is placed in a cell for delivery to the reading station or file. The recording and development process takes about three minutes.

In recording, a concentrated beam of electrons records data on 1.3 by 2.7 in. silver halide film chips. The data is recorded on the film in the form of light or dark spots "painted" by an electron beam at the rate of one-half million bits per second. Nearly five million bits of information can be packed on a single film chip. The film chips are housed in small plastic cartridges, or cells, containing 32 chips - 150 million data bits (information equivalent to that in three typical encyclopedia volumes). Cells are retrieved from trays resembling egg crates by a pneumatic system which delivers them to recording or reading stations.

When data is to be read, the selected chip is positioned before a cathode-ray tube flying spot scanner. Data is read by the scanner at a rate of approximately 2.5 million bits per second. Electronic control units built into the photodigital system link it to a computer and regulate operation.

Built for the Atomic Energy Commission, the only existing unit has been installed at the University of California's Lawrence Radiation Laboratory. IBM does not intend to to market this device, however, since a market review has indicated that it could not be done profitably.*

6.3.3.1.2 Precision Instrument UNICON System. - The UNICON (Unidensity Coherent Light Data Processing) system utilizes coherent laser radiation in a rotating optical system to create and detect (record and reproduce) information elements in two dimensions, through diffraction - limited evaporation of a special Unidensity recording medium. An information bit is represented by a diffraction limited "hole"

*UACSC Trip Report (Contact No. 30), IBM Corporation, San Jose, California within the Unidensity film layer. It results from the evaporation of the Unidensity medium by means of imaging the aperture of a laser to a three-dimensional ellipsoid of revolution (Debye ellipsoid) as shown in Figure 35. Using this technique, recorded information bits (holes) with diameters on the order of one micron have been produced. Instantaneous readout during recording occurs by detecting the diffracted laser radiation during evaporation of the Unidensity medium. Subsequent readouts are accomplished at drastically reduced laser power to avoid further destruction of the Unidensity film layer.

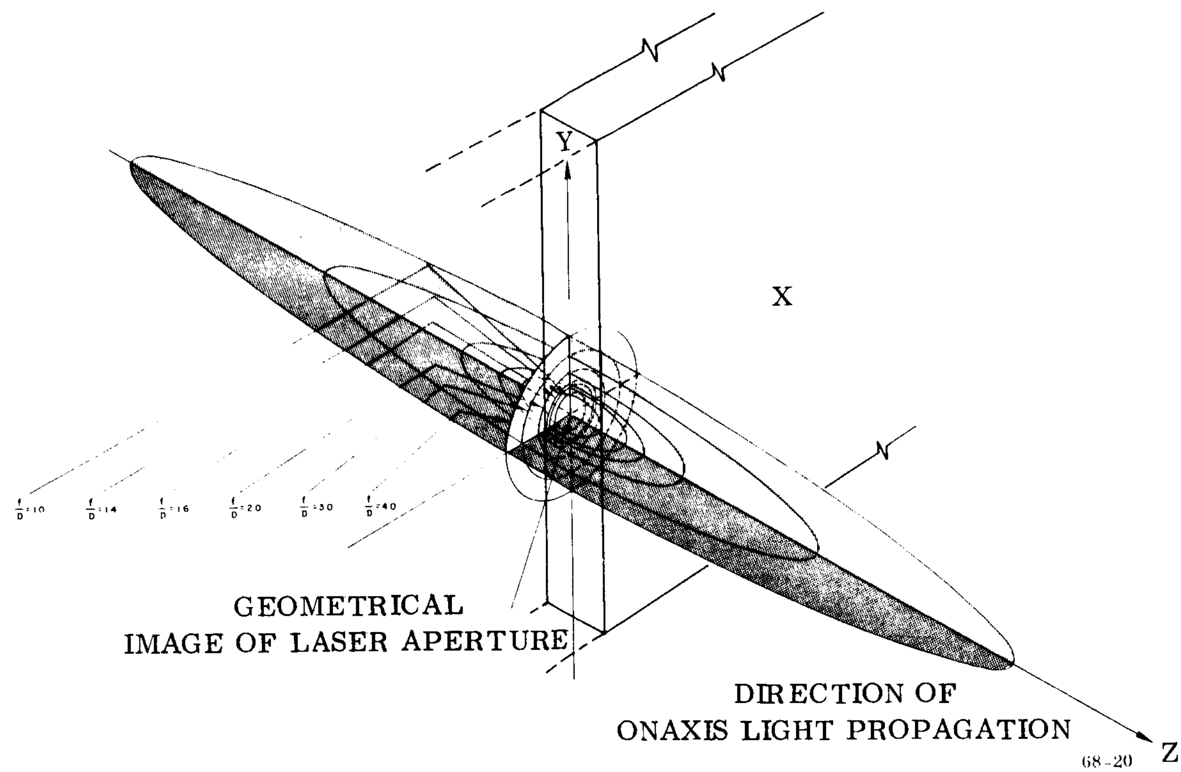


Figure 35 Debye Ellipsoid Principles

An operating functional prototype of the UNICON system, UNICON-6, is capable of 6 mHz record/read rates, and has an information density of 6 x 10^8 bits per square inch at 1 cm/second medium velocity. The medium is carried by a 16 mm film base of 100 feet length, giving 50 minutes continuous recording/reporducing time. Further UNICON development is being directed towards 10, 20 and 30 mHz bandwidth. Maximum performance is expected to extend towards a 100 mHz UNICON Unidensity Coherent Light Data Processing system.

In order to establish a two-dimensional pattern of information bits, the Unidensity layer is helically carried at slow speed (1 cm sec) around the imaging circle of the laser aperture, which is formed by a diffraction-limited objective rotating with high speed around the helical axis. The velocity of the laser image is 18.6 meters/ sec at 1800 rpm of the rotating optical system. The inclination of the image azimuth against the transport direction of the Unidensity film is 1 deg, 32 minutes. This produces helical unit records of one line length containing 600,000 bits each, with spacing between individual unit records of 2 bit diameters. The width of the information-carrying area of the 16 mm Unidensity film is 8 mm. Total capacity of one UNICON Memory System is 88 x 10^9 bits for a 16 mm Unidensity film real of 100 feet.

While the UNICON-6 uses a 16 mm wide recording medium, a 35 mm width of recording medium would be more practical, from the standpoint of reducing access time. The track layout would be a illustrated in Figure 36. The 35 mm width in this figure includes 4 mm in the upper edge track for a binary file accession number, 1 mm in the lower edge track for a time track and a 30 mm wide track for recording of data. Data tracks 1 meter in length would be laid transversely along this width, each track providing a file length of 10^6 bits. The track separation nromal to the direction of recording would be 4 microns. Due to the low incidence angle, this gives a 167 micrometer track separation distance along the length of the recording medium. This configuration would provide a total information capacity of 10^{12} bits in 528 lineal feet of recording information:

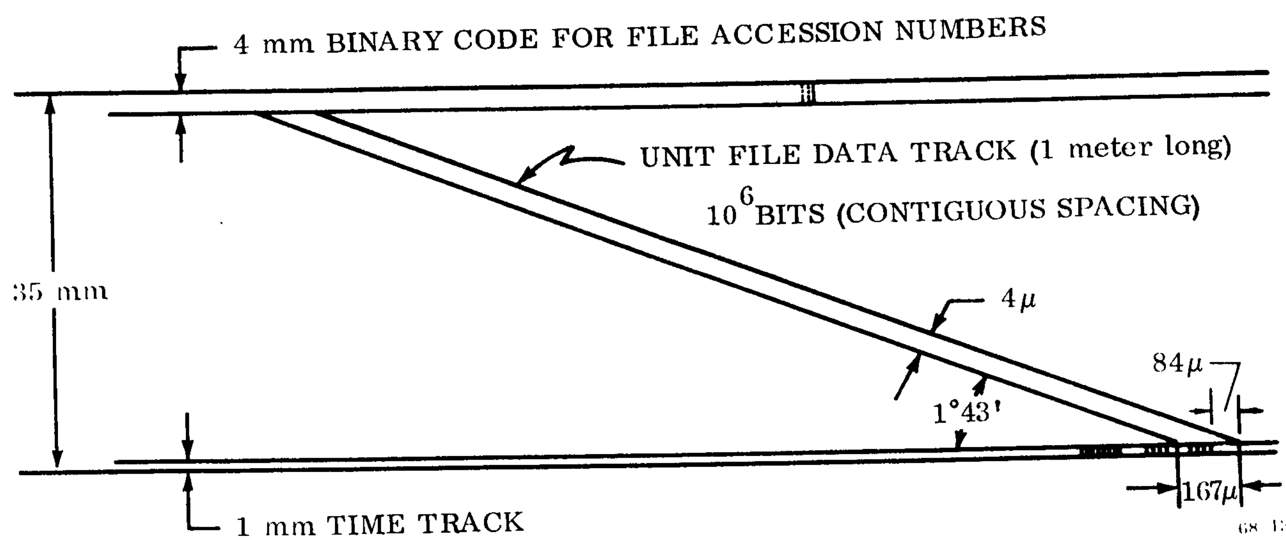


Figure 36 Track Layout

In order to synchronize the high speed optical rotary system and the low speed transport of the recording medium with the real time of the signal, three separate servos are combined to a complete servo system. The time base for the servo control system is provided by laser prerecording on the medium, thus yielding a separate control track. Readout of the servo track during recording/reproducing utilizes another laser of low power.

Being a strictly optical and quantum-mechanical data process, mechanical contact is eliminated between writing/reading means and the recording/reproducing medium. This is quite different from magnetic data processing, which is a "classical" data process essentially determined by the gap width of the writing/reading head and its length. Entirely different from erasable magnetic data processing, the UNICON system represents a nondestructive documentary type of data storage and retrieval. Eliminating photographic development, UNICON coherent light recording provides a permanent record which is instantaneously reproducible during recording as well as during unlimited nondestructive secondary readout.

Limited only by the wavelength and power of the record/reproduce laser, the principle of the UNICON system provides up to two orders of magnitude higher performance in respect to the two critical parameters of information density and frequency bandwidth, compared to present capabilities using magnetic recording techniques. When perfected to the operating production model stage, the UNICON concept would seem to be ideal for the type of archival storage and retrieval (digital) required by the AAPA.

6.3.3.2 Image Storage - For image storage, an assortment of microfilm techniques are available. Although the traditional microfilm on reels is by no means obsolete, there has been increasing use of microfilm in the form of fiche, which contain several rows of images (corresponding to the length of a typical research report) and which can be handled and interfiled easily. Also important has been the use of aperture cards, which are standard punch cards for data processing machines in which have been embedded one or more microfilm frames. Aperture cards are key-punched with index codes describing the record they carry in microfilm; when the aperture card is selected in a search, the desired record is immediately available. Aperture cards are particularly attractive as a means of storing, retrieving and duplicating engineering drawings.

The saving of storage space is an obvious objective of the use of microfilm, but it is one that should be approached with caution. Extremely high reductions are possible, but they are rarely economical because of the penalties in higher filming expense, higher projecting expense and the difficulty of handling tiny images. A recently adopted Federal standard for microfiche reflects the view that a reduction ratio of about 1:20 is the most economical with present technology. A record reduced by this factor takes up one four-hundredth of its original area. Probably even more important than space saving to many users at present is the fact that records on microfilm can be duplicated more cheaply than the full size originals.

It would appear at the outset that some of these various microfilm formats are not particularly well suited for the AAPA. Aperture cards employ a unit record format, and there are few unit records anticipated for the AAPA. Microfiche, since it is intended to correspond roughly to the length of an average research report, might be of some use in storing copies of the AAPA experimenters' published results;

however, it would be misapplied if used to archive the bulk image data which will constitute the data base. When the bulk record "batches" of the AAP experiments data are considered (viz., all the photographs or plot data from a specific instrument), these approaches appear generally lacking. The two techniques considered below appear better suited for this type of bulk image records.

6.3.3.2.1 Micro-Vue - Micro-Vue, a development of the Republic Aviation Division of Fairchild Hiller, employs a "microfilming of microfilm" technique which is somewhat analagous to the PCMI process described previously. The information storage medium which was developed for the Micro-Vue is a laminated photographic film chip 4 x 5 in. in size. The information images are arranged on the chip in a 99 x 99 matrix, inside a useable area of 3.2 x 4.3 in. This size matrix can contain 9801 frames, or pages, of information. Matrices of any desired size can be prepared to increase or decrease the amount of information stored with proportionate change in the size of the Micro-Vue equipment. Linear reduction of up to 300:1 from the original material is accomplished through microphotographic techniques. Final resolution is in excess of 500 lp mm. The photographic negative master is utilized to prepare the desired quantity of contact print positives which are then laminated and otherwise processed for dissemination and use in the Micro-Vue. Film processing requires controlled conditions to preclude the incorporation of foreign materials, dirt, and lint, up to and through the process of lamination. Once laminated, film chip image quality does not suffer through handling and casual use. Scratches and dust are not of particular concern since the Micro-Vue optics are focused on the film emulsion inside the laminate.

For retrieval, the compound table used to hold the film chip is positioned in the appropriate row and column coordinates by lead screws. Motion in both axes accomplished by small stepper motors. Film chips may be changed manually or by use of a 10 or 20 chip automatic loader. Any of 200,000 pages may be addressed within 30 sec. Access to the desired frame of information may be accomplished either electronically by digiswitch or by slew switch activation.

Figure 37 shows the basic two-stage Micro-Vue processing operation. The output of the final camera is a glass-plate negative containing all 10,000 frames. From this negative, the required positive copies are made by contact printing. These copies are on flexible film and are laminated in a protective plastic covering before distribution.

The archieval storage of records, encyclopedic storage of reference information such as abstracts, books, tables and charts, dissemination of professional periodic literature and reports, and similar record-type material with little or no requirement for revision are natural subjects for Micro-Vue treatment. It is interesting to note

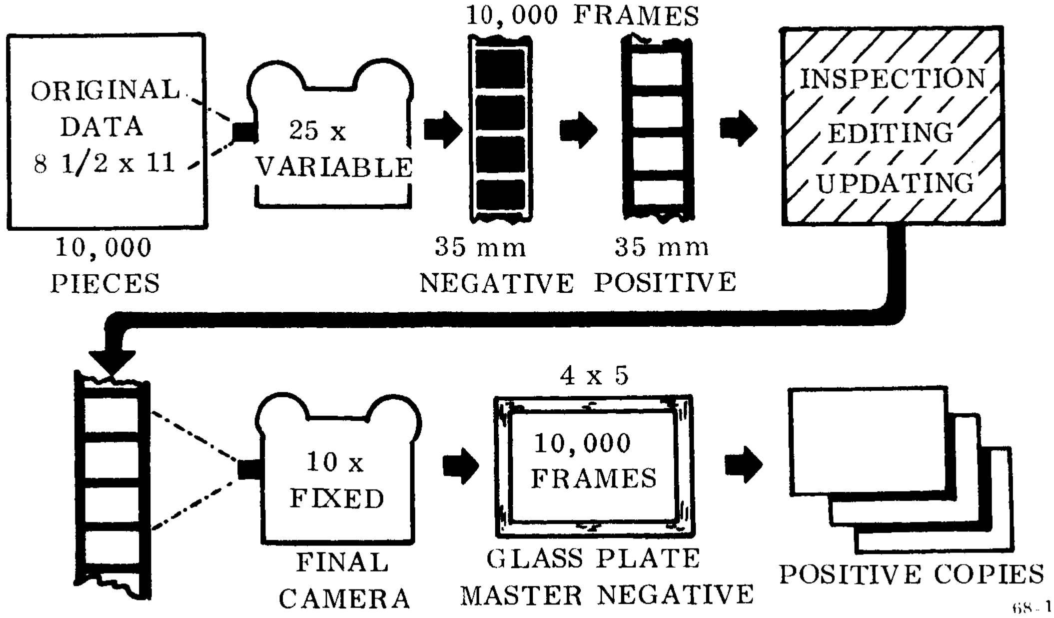


Figure 37 Basic Micro-Vue Production Scheme

the fundamental similarity between Micro-Vue and basic microfiche, the principal difference being the size of the record groupings (10,000 for Micro-Vue; nominally 30 for microfiche). This feature would make the employment of Micro-Vue for the volume image-record "batching" required of the AAPA seem quite attractive. A limiting drawback, however, is the requirement for specialized viewing equipment at the individual user sites.

6.3.3.2.2 Roll Film - Roll film storage offers the advantages of microfilm, but like magnetic tape, requires serial searching to obtain access to the recorded information. Attempts to shorten the search time have concentrated on electronic means of scanning and, in some cases, on using shorter length rolls of film. In some systems, the index information is recorded in digital form as small clear and dark areas beside the document image area. Consequently, in roll storage, the basic intent has been to increase search speed by reducing the length of film to be searched, by increasing the searching speed, or by a combination of both.

One apparent disadvantage of roll film storage is that it forces the use of information in a fixed physical arrangement. File updating and reorganization can only proceed by extensive splicing or by complete recopying. In many applications, these restrictions cannot be endured even though mechanization is utilized to increase searching speed. For such situations a unit record concept is usually employed. However, due to the ease of manipulating roll film in contrast to unit records, mechanization of roll film storage can be attained in many cases at lower cost. Under such conditions, roll storage should be considered if any possibility exists that its relative file inflexibility can be endured. The AAPA represents just such an application.

Recordak Miracode - Microfilm Information Retrieval Access Code (Micracode) is essentially standard roll microfilm with digitally coded optical indexing. The basic access unit of the system is a 4 x 4 x 1 inch magazine containing 100 feet of 16 mm roll film. Both the document images and the digital code are recorded on the film.

The indexing system is extremely versatile and open-ended. Every document field on the film, which may consist of one or more documents, is assigned an identifying code field composed of one or more code columns. The columns employ an arrangement of clear and opaque rectangles as bit notations. By the use of 12 bits, each column is afforded BCD representation of the numbers 000 through 999. A utility bit and parity bit bring the column total to 14 bits. These code columns can be configured in virtually any combination to depict more than a million retrieval terms, provide fixed or open code fields, and apply machine logic search capabilities. The available machine logic modes consist of the following operations:

Word "OR" - $A + B + C + D$

Phrase "OR" - $(A \bullet B) + (C \bullet D)$

Word/Phrase "OR" - $A + (B \bullet C \bullet D)$

"BUT NOT" - $A \bullet B \bullet C \bullet D$

Identity - $(=)$

Boundary Conditions - $(\geq, \leq)$

Retrieval is accomplished by random serial access. It requires that the appropriate search criteria be keyed into the Lodestar Reader Printer which can scan the entire 100 foot roll of film in 10 sec. The first page of specific coded document which answers the search request is located and is automatically displayed, ready for viewing or printout. Other documents, answering the same parameters, may be located by merely depressing the search button again. Using this mode of operation, a file of 10,000 pages randomly contained in five magazines can be searched in less than 2.5 minutes.

A basic Miracode system consists of the following items:

Miracode Microfilmer - This planetary 16 mm unit, which includes convenient operator controls, is especially designed for high-resolution recording and accurate illumination adjustment.

Miracode Retrieval Station - This equipment can store, within fingertip reach, approximately 480 film magazines. The station is modular in design and can be arranged in several configurations. It contains all the controls necessary to initiate and complete a search cycle and utilizes the Recordak Reader Printer which can search the binary code on the microfilm at high speeds. Selective prints of documents in several print modes can also be provided by this unit.

Assuming 1000 document pages per 100 foot roll*, it can be seen that a single retrieval station can service approximately one half million pages. Or, for those inclined to think of storage in terms of bits, the conversion can be made as follows:

Nominal resolution employed: 150 lp mm = 3800 lpi

Normal reduction: 24:1

Resolution preserved from original: 158 lpi

Total bits preserved from original (8 1/2 in. x 11 in.): $8\ 1/2 \times 11 \times (158)^2 = 2.34 \times 10^6$ bits/page

Total number of pages: 1000 (pages per 100 feet roll) x 480 rolls = 480,000 pages

Total bits per station: $(2.34 \times 10^6)(4.8 \times 10^5) = 1.12 \times 10^{12}$ bits

Which perhaps shows that it is frequently misleading to make such conversions and then be awed by the apparent volume of "data" which results. In many cases, such as this one, the more significant figure is the number of records - which was shown to be nearly one half million. The fact that each record has a "record length" of more than a million bits does not significantly complicate the overall problem of record storage and retrieval.

The approximate component costs for a basic Miracode system are given below.

Initial Costs:

16 mm Planetary camera	$12,000
Film Processor (Processing time - 30 min/100 ft.)	4,000
Retrieval Station	18,000
Total	$34,000

* A nominal estimate

Operating Costs:

Film (150-200 lp mm), per 100 ft. roll	$3.50
Development (per roll)	.25
Magazine	1.25
Hard copy printout (each)	.08

6.3.3.3 Digital/Image Conversion - Since the AAPA data base will consist of both digital and image data, it is worthwile to review some of the techniques commonly employed to cross-convert these data forms. Image to digital conversion may be accomplished by a variety of optical scanning methods. One of the most common applications of this technique is the facsimile transmission of image data. Although it has been determined that this type of data dissemination will not be required for the initial version of the AAPA, image to digital conversion should be considered if automatic processing techniques are to be applied to the photographic and graphical portions of the data base.

The possible requirement for converting digital data to image data has a somewhat greater likelihood. One application might be the preparation of surrogate data for dissemination to users. This data could consist of both alphanumeric and digital line data, and might be prepared in response to single requests or for mass distribution. Some of the devices commonly employed to accomplish both image-to-digital and digital-to-image conversion of data are discussed below.

6.3.3.3.1 Image-to-Digital Conversion

The various storage media which are being considered for the data archives have been discussed previously in this section. Normal cost effectiveness considerations indicated that the various magnetic devices may be most suitable for the storage of digital data, while high resolution film media are best suited for most of the pictoral and graphical information.

Nevertheless, there are applications where it is desired to convert non-digital data, particularly graphical data, to a digital format for convenience of computer processing. For this reason, several of the optical scanning and digitizing, and subsequent output plotter line of equipment have been reviewed. The following paragraphs are included to present a partial but current view of this technology as it pertains to the AAP Data Archives.

The use of plotting equipment for conversion of line path and position data from maps, graphic charts, and similar two dimensional copy has been found desirable for several applications. Equipment is presently available for this conversion in a wide range of speeds, accuracies, and capabilities for copy handling. It is desired to determine costs, availability, and limitations of present equipment for performing the required functions. Survey results are presented in the following sections of this section.

Ten manufacturers of plotting and digitizing equipment have been reviewed. Verbal reports indicate that several additional sources are possible; however, at present no data, no sales literature, or other recorded substantiation is available. Until such is received, therefore, these equipments will not be included in the survey. Those which have been covered in this survey represent a wide range of size and complexity, performance, and cost.

The major features are summarized in Table XVIII. While these may be divided into categories in several ways, Table XVIII used copy size as the criterion for dividing the equipments into three categories, small, medium, and large. As may be noted under 'Remarks', several equipments have options which affect price, therefore direct absolute comparisions between equipment proces are not possible. For three units, Calma, Itek, and Wayne-George, quotations were requested with specific output interface configurations. In addition to the market survey, several basic approaches to position sensing and recording have been explored.

In principle, there are three basic methods available for determining location of a spot in relation to a fixed reference. The majority of the equipments studied utilize the most straightforward method X - Y coordinates to designate position. However, there is a great degree of variety represented by various equipments in the technique for X - Y determination. A second method, used in only one equipment, is specifying a point as RHO/THETA, that is, a radius and an angle. This appears to have a limited size capability due to required rigidity of the radius arm to maintain satisfactory accurady. This limitation is reflected in the maximum capacity of 7 x 7 in. for the Grafacon equipment. A third method, which admittedly has some disadvantages, is in the use of two fixed length arms (or radii) and the measurement of two angles. No present equipment utilized this approach.

A fourth possibility, which is not truly a tracing technique, is in the use of a limited small area scanning, with the area location controlled manually. In this concept, gross location would be selected by the operator and fine scan of this area would then digitize the coordinates within this area.

In the X - Y coordinate measurement method, the techniques fall into two primary classes:

1) X - Y determined by rotary encoders driven by a linear motion of the tracer stylus, with X - Y components resolved by the mechanism.

2) X - Y determined by proximity sensing using a special stylus, with capacitive or similar coupling.

TABLE XVIII

FILE CONVERSION DEVICES

MFR. and Model No.	Size Copy (in.)	Coordinates	Resolution (in.)	Tracing Speed (in./sec)	No. of Bits	Accuracy (in.)	Table Size (in.)	Elect Size (in.)	Total Weight (lb)	Price ($)	Interface Features	Special Features	Miscellaneous Remarks
SMALL													
Sylvania (DT-)	11x11	X-Y	~.003	0.55	12/12	0.11			50	7000	None yet	Capacitive stylus, can be used with CRT display	See Nonker's memo 4/19 - parallel problems with projection
Data Equipment Co. Grafacon Mod 1010A	10.24x10.24	X-Y	.01	45	10/10	0.02	20.125x 24.125x 1.5	12x9x19	42	9600		Capacitive stylus, error sense for more than 1 increment/step	Discrepancy in size and weight between 2 data sheets
Data Equipment Co. Grafacon Rho-Theta	8x5.75 or 9x5 or 7x7	R-Θ	~	~	~	0.01	No table included	est. 10x6x4		1185	Analog output only		Requires digital conversion for output
UACSC Proposed Design	8x5.75 or 9x5 or 7x7	R-Θ	.003			est. .003/ .005	est. 10x24x6	21x7x15		est. 7700/6145 (watts) 20/40 hr (labor)	Serial Mod 33 ASR TTY		
MEDIUM													
Auto-Trol 3400	24x36	X-Y	.001 to 2.0 (selectable increments)	75	17/17	.004 (overall) .001 (repeatable)	est. 48x36x48	22x28x30 (recorder)			Incr. tape recorder 500/sec 200/556 bpi	Run ident. sequential numbering	Options; mag tape, paper tape, punched cards, projection capability
Wayne-George T-36	36x36	X-Y	.001	1.2	16/16	.002 (RMS)	57x48x31.5	est. 19x16x9		11,000	ASR 33 TTY	Error check each 4 in. (X, Y) Parameter entry via switches	Options; mag tape, punched cards
Data Tech. Inc. IPD(CDT)	36x36	X-Y	.0005/.001	50	17/17 (for max. resol.)	.001/.002	est. 48x40x28	est. 19x16x11	300	6000/3000	Paper tape		Price dependent on options Options: manual data entry switches, projection capability
Itek PCR-2										17,000			
Benson-Lehner 099D	16 wide est. 14 to 30 length	X-Y					est. 48x30x40	est. 20x22x36			Optional paper tape, mag tape, cards		Primarily for strip chart data reduction
Calma - 303	18x24	ΔX-ΔY	.01-.15 (selectable)	2 to 30 (depend on resol)	6/6	.012	58.5x30x34	22x28x30 (recorder)	400	22,500	Incr. tape recorder 500/sec entry 556 bpi	Keyboard control and data entry	Options: film projection, typewriter data input, paper tape, punched card output, direct input to IBM 1130, Calcomp plotter read back
Concord Controls FAGR	23x32	ΔX-ΔY	.01	1	4/4	.01	48x30x42	48x18x42	1500		Direct to PDP-7		Magnifying optics considered determinant

TABLE XVIII

FILE CONVERSION DEVICES (Continued)

MFR. and Model No.	Size Copy (in.)	Coordinates	Resolution (in.)	Tracing Speed (in./sec)	No. of Bits	Accuracy (in.)	Table Size (in.)	Elect Size (in.)	Total Weight (lb)	Price ($)	Interface Features	Special Features	Miscellaneous Remarks
LARGE													
Itek PCR-1	40.5x40.5	X-Y	.003			.05 (overall) .03 (repeatable)	90x71x36			85,000	IBM 526 card punch		Joystick positioning, multiple magnifying heads
Calma 480	48x60	ΔX-ΔY	.005	15.6 at max. resol.	6/6 (Δmode)		76x75x55	22x28x30 (recorder)		29,000	Incr. tape recorder 500/sec 556 bpi	Keyboard control and data entry, selectable or variable increment modes	Options: typewriter input, paper tape, punched card output, or direct to IBM 1130, Calcomp plotter read back
Wayne-George T60	30x60	X-Y	.001	1.2	17/17	.002 (RMS)	81.5x52x31.5	est. 19x16x9			ASR 33 TTY	Error check each 4 in. Parameter entry via switches	Options: mag tape, punched cards
Data Tech. Inc. IPD 60	30x60	X-Y	.0005/.001	50	17/17	.001/.002		est. 19x16x11		6000/3000	Paper tape		Price dependent on options Options: manual data entry switches, projection capability
Benson-Lehner LARR-VH	48x60 (handles roll length)	X-Y				.010 point-to-point		est. 20x22x36			Optional printer, paper tape, mag tape, punch cards		
Benson-Lehner LARR-D	48x51	X-Y	.004	140				est. 20x22x36			Optional cards paper or mag tape	Copy mounted on drum for gross Y position - allows seated operating position	
Gerber GLACD	48x60	X-Y	.001			.005					Optional paper tape, punched cards, typewriter		

In rotary encoder sensing, a further subdivision into absolute and incremental types is possible. Both of these classes are represented by equipment in Table I.

Prior to a full discussion and comparison of these tracing techniques, it is worthwhile considering some idealized characteristics of a general tracing system. While these characteristics may vary dependent on use, (e.g., tracing of patterns and regular geometric figures may require different characteristics than the tracing of irregular control lines of weather maps), an optimum system, if not resulting in excessive costs, would be universal, including features to cover all applications.

Characteristics of such a system are presented here, based on a limited experience with tracing operations, and in some cases based on intuition.

1) Copy Size: It would appear that copy sizes of up to 48 x 48 in. should be accommodated by a general tracer. In addition, it would be advantageous to have a capability for roll size copy.

2) Resolution: This should be selectable within some limits due to various possible applications. A lower limit, considering manual control, would seem to be around .001 in.

3) Accuracy: Accuracy should be compatible with resolution or better. A low limit of .001 in., equivalent to the resolution limit, would seem to be adequate.

4) Digital Format: The use of absolute X - Y (or other) coordinate position would seem to be required as a reference for any system. The use of incremental steps within specified limits has advantages in minimizing storage requirements and very probably it is advantageous in many computer processing procedures (i.e., the incremental provides a vector representation, saving a computation of this value).

5) Digital Output Interface: For flexibility of use, an ability to interface easily with mag tap, paper tape, punched cards, and directly to a computer is desirable.

6) Stylus Force: This should be minimum to minimize operator fatigue. However, when working with a tilted or vertical surface, some braking or counterbalancing is required to avoid drift when released.

7) Stylus Inertia: While this should be reasonable low to allow easy change of direction for irregular contours or abrupt angles, it would seem that some inertia to eliminate 'shaky' hands and vibration effects is necessary.

8) Type of Copy: To cover a maximum of applications, the tracer should be capable or working with transparent copy, opaque copy, or projected images.

In addition to these specific characteristics, other desirable features of a tracing system may be considered:

- Recording ability. In the simplest case, this may be pencil writing on transparent overlay over the material being traced. This would seem to be very desirable for complex copy, so that operator will know that sections have been completed.

- Repeat back. This is desirable to verify operator's tracing accuracy. Recorded on transparent medium, the play-back copy superimposed on the traced copy would verify tracing accuracy.

- Coding of enclosed areas. An ability to close a loop (or contour line) and indicate by coding the significance (gray shade, color, etc.) of the closed area would seem desirable for some applications.

- Ability to follow standard templates, (e.g., straight lines, circles, etc.) As an alternate to this, manual encoding of these and standard symbols would seem desirable.

- Correction capability. A means for deleting errors and inserting new data is desirable.

- The ability to code lines or areas for color or shades of gray would be useful for certain applications. In addition, coding for dashed or dotted lines might be useful.

Based on these assumed characteristics, the equipment of Table XVIII may now be discussed.

1) Small size class. - Rho-theta - The rho-theta tracer is small, light and portable, and by far cheaper than others. However, as noted, additional circuits are required for conversion to digital form for storage. With this system, position resolution is variable with radius. For a 10 in. radius, a 10' angular resolution results in a 0.029" position resolution. For shorter radii, the resolution improves, as does accuracy. An encoder with 1' or less resolution is available, therefore a system meeting a .002" to .003" resolution is practical. Parts costs and labor estimates are presented for this design, including output registers and circuits for interfacing with an ASR 33 TTY.

Electronic Tablets - The Grafacon Mod-1010 and Sylvania DT-1 both use electronic sensing of the stylus position. In principle, the two are similar; both use a horizontal/vertical printed circuit grid carrying signal voltages, and a capacitively-coupled probe as the stylus. The Grafacon uses a coded bit stream which is unique for each of the horizontal and vertical grid lines. Recognition of the pattern received by the stylus using a decoder determines the X - Y position. The Sylvania tablet uses phase-sensing for X - Y determination. The stylus for each of these is light-weight and easily movable. It is possible that the resolution limit of the Grafacon (.01") is imposed by spacing requirements for the signal voltages used. The Sylvania may be less limited because grid line to grid line voltage gradient is small due to small phase difference. The same reason may be applied to the stated accuracies of the two units. Although the price of the Grafacon is considerable higher, some useful error-checking circuits are included. The Sylvania unit is claimed to have a three dimensional capability, (i.e., stylus height above the table allows a Z digitization of three bits). No description is provided on how this is achieved.

2) Medium Size Class. - Several equipments in this class are duplicated in the large-size class. All X - Y devices in both classes are basically the same, having optical or magnetic rotary encoders. Each has options on output interface, and most have special features which probably could be duplicated by others if required. Two major differences within the class are the Calma and Concord Control units which provide incremental output data rather than absolute position. The Calma Mod 303 was viewed at the Calma New England field office. No serious problems were apparent. The flexibility of scaling, calibrating, and location of reference points appears advantageous for many digitizing applications; however, the procedures are sufficiently complex that operator training would be required to perform some setups efficiently. The feel of the stylus was adequate and irregular lines were traced with reasonable accuracy. A Calcomp Plotter was used as repeat back for the demonstration unit and required manual pen-up, pen-down operations for some tracing operations. Aside from speed limitations when using the Calcomp Plotter, the entire system operation seemed satisfactory.

According to Calma literature, the 480 has an additional unique feature of a variable incremental digitizing depending on stylus velocity. It may be reasoned that where an operator has confidence in tracing, such as straight or gently curved lines, and therefore traces relatively high speed, the need of data points with high resolution is not required. As the operator finds it necessary to slow down, to negotiate highly irregular lines, digitizing at a corresponding high resolution becomes desirable. Note that the Model 480 has controls for either fixed resolution or variable resolution digitizing. The main advantages of the variable mode are in allowing higher tracing rates without overloading the recording device, and more efficient utilization of computer time in processing the recorded data.

An installation of the Wayne-George T-36 tracing table at the Physics Dept. of the University of Mass. was reviewed. The use of this table for determining particle paths in a bubble chamber for nuclear particle studies requires the use of a film protection, and digitizing of only a few points on the track. The setup of the W/G table for this application is probably not typical of many contemplated users. A projection system mounted above the table projects the film image onto a white formica table top. The cursor is manually positioned to the desired points, and a reading obtained and recorded by actuation of a footswitch. The associated card punch is adequate for the point digitizing application, but would severely limit speed for line digitizing. The feel of the cursor in the following lines was better than the Calma, however, part of this may be due to the fact that lines traced were regular curves, (circles, spirals, etc.) as compared to irregular contour traced on the Calma. Only one difficulty has been encountered in repeatability, apparently due to slippage of the belt on the encoding pulley, and this was corrected by parts replacement and adjustment. There is no provision for back lighting from below the table.

3) Large Size Class. - With two exceptions, the general comments regarding the medium class apply to this calss. The exceptions are the Itek PCR-1 and the Bensor-Lehner LARR-D. Considerable difficulty has been encountered in obtaining the Itek PCR-1 price. Itek stated that the delays were caused by attempts to reduce the original price. In discussions, it was disclosed that a new design may be in process. Part of the high cost of the PCR-1 is probably attributable to the power positioning of the cursor from a joy-stick. Also, magnifying lens systems are used on this equipment, which probably add to cost. Optical magnifiers are also available for Bensor-Lehner equipment. Bensor-Lehner lists several digitizing equipments in addition to those shown in Table I. Those specifically used for strip-chart analog record digitizing and one specifically used only for film projection digitizing were excluded from this report. The LARR-D has a unique feature, in that a drum is used as the copy mounting surface. While not stated in detail, it is implied that the drum is used for gross Y position, so that an operator may remain seated in a fixed position and have access to top and bottom of the copy by rotating the drum. It it understood that small Y increments are sensed by the cursor, along with all increments. The method of composing total Y from Y drum + Y fine is not described. If all Y position is derived from drum rotation, it is visualized that a high degree of operator coordination in negotiating lines at other than horizontal or vertical would be required.

Gerber has several types of accessories for their Model 32 and 75 plotting tables. However, since these are accessories only, the pricing is dependent on availability of the plotter, therefore the only item listed in Table I is the Large Area Coordinate Digitizer.

Additional information on price and performance for blank items in Table XVIII has been requested.

A second area of effort on this program has been the investigation of the semipantograph as a location sensing device. As noted previously, some disadvantages are apparent in the use of this technique, and others not apparent may be the cause of no vendors utilizing this approach. There are two features which would seem to make the system attractive.

a) Avoids the requirement for linear to rotary conversion. Also, multiple revolutions of the encoder are not required.

b) Curved line stylus motions are more readily accomplished.

Without detailed design, it is not safe to assume that less stylus force is required than for equivalent X - Y systems. However, the fact that only two rotary bearings are required, and the allowability of light weight carrying arms would tend toward light moving forces. Also, the relative cost of this system, using simpler mechanisms, should be less than equivalent X - Y systems.

One major disadvantage of the system is in varying resolution and accuracy over the table area. Preliminary analysis has shown that X and Y variations tend to offset each other, helping to make absolute position less inaccurate as a function of position than would appear at first glance.

A second major advantage is that a calculation is needed to convert α/β to X - Y. Originally, it appeared that the conversion was fairly complex. By use of a rotary housing as reference for the second angle (β) sensing, this has been simplified considerably. It should be emphasized that additional analysis is required to obtain performance characteristics.

6.3.3.3.2 Digital to Image Conversion - Stromberg Carlson SC - 4020/4400/4060 employs their CHARACTRON® Shaped Beam Tube to convert data stored in digital form to a CRT image. This image can then be transferred to microfilm or photorecording paper. In the 4020 Computer Recorder, the CHARACTRON® can be used to generate characters, symbols, lines, or curves for output to standard 16 or 35 mm film and/or photorecording paper. The 4400 Document Recorder generates characters, symbols and also the visual codes required by semiautomatic and automatic storage and retrieval systems. The 4060 Data Recording System, the most versatile of the three, contains its own sixteen-bit general purpose computer. It has the same capabilities as the 4020 and 4400, plus a host of others. It also provides a built in system which processes the exposed microfilm automatically.

3M Electron Beam Recorder (EBR) System is designed to take computer-generated digital information stored on magnetic tape, convert it to image form, and output it as a fully developed microfilm image - all in one continuous operation. The microfilm imaging is accomplished by electron beam recording techniques, using a fixed character-set format. The fully developed microfilm output is accomplished using 3M's heat-developable 16 mm Dry-Silver microfilm.

6.3.4 Surface Deformation - General Electric's Thermoplastic Recording Process presently in the product development stage, portends to combine the advantages of photographic and magnetic recording.

Thermoplastic tape is a "sandwich" of material consisting of a base layer on Cronar, an extremely thin transparent layer of electrically conductive material, and a thin layer of a thermoplastic. This tape is manipulated as is magnetic tape but can be used at much greater storage densities. Using a highly refined electron gun, digital information has been recorded at a density of 80 million bits per square inch.

In recording, information is represented by a charge pattern which is applied to the tape by an electron gun. The tape with the charge image then moves over a heater which heats the thermoplastic sufficiently to allow it to deform under the electrostatic attractive forces between the charge on the surface and the conducting layer of the tape which is held at a uniform potential. The original charge pattern is therefore transformed into a groove pattern where the depth of the groove is at any point proportional to the local charge density and hence to the original signal strength.

When the grooves are formed, the information is available for readout or the tape can be wound on a take up reel for future playback. The complete process from charge delivery to the creation of readable information can be achieved in milliseconds. At this point in the recording cycle, modification can be made such that the information can be made available in real time for use in displays.

At this point the information exists on the tape in the form of depth modulated grooves. For meaningful playback, Schlieren optics are used to generate an intensity-modulated image to which the eye or photosensors are responsive. In this technique, two sets of stops, called Schlieren bars, are added to the ordinary projector configuration of a condensing lens and a projection lens. When used with thermoplastic recording, the projection optics allow a gray scale to be achieved from the original microscopic grooves on the tape.

With these characteristics, it is evident that the thermoplastic tape combines many advantages of both photographic and magnetic recording. It has the resolution capability for image formation and an immediate access to information and capability for erasure and re-use as with magnetic tape. Current efforts with this technique are concerned with investigating the implications encountered when considering its use in analog recording, digital data storage, display systems and electronic photography.

In a recent Air Force application, a thermoplastic recording system has been developed to record image data having 8000 by 5000 picture elements on a 3 by 2 in. area of tape, while measuring sixteen shades of gray.

6.3.5 Review of Applicable State-of-the-Art ADP Main Frame Hardware - At present, the data storage and retrieval requirements of the AAPA are not defined to a degree of completeness which will permit selection of specific items of computer hardware. Tables XIX and XX represent the broad range of main frame equipment which seem suitable to the AAPA environment. The rough division into scientific and business oriented devices facilitates one useful type of overview. If the AAPA is to function principally as a straightforward information retrieval system, the business-oriented computers seem operationally adequate. If the archive operation develops more toward retrieval and specialized output processing to suit individual users' needs, it can be generally assumed that this will constitute an additional need for scientific processing capability. In this case, the procurement of main frame equipment more oriented towards scientific processing may be justified.

TABLE XIX

PERFORMANCE PARAMETERS FOR SELECTED BUSINESS-ORIENTED COMPUTERS

Entry No.	Company	Model No.	Date of 1st Installation	Bits per Micro-Second	Bits Per Cycle	Central Processor: Cycle Time (μs)	Central Processor: Number of Words	Maximum Bit Capacity	Rank By Maximum Bit Capacity	Word Length: Bits/ Character	Word Length: Bits/ Word	Cost Per Bit for Maximum Memory Capacity	Cost for Maximum Memory Capacity	Remarks
1	Burroughs	8500(1)	---	384.00	192	.5	262K	12,576K	1	6	48			Thin Film Memory
2	Honeywell	200/8200	4/67	64.00	48	.75	262 to 1048K CHAR.	6,288K	2	6	48	.374	$2,350,050	
3	Honeywell	200/4200	12/67	42.75	32	.75	65 to 524K CHAR.	3,144K	5	6	24	.353	1,111,325	
4	RCA	70/55	7/66	38.00	32	.84	65 to 524K CHAR.	4,192K	3	8	32	.261	1,093,700	
5	Honeywell	200/2200	12/65	32.00	32	1.0	16 to 262K CHAR.	1,572K	8	6	--	.316	496,800	
6	IBM	360/44(3)	12/66	32.00	32	1.0	8 to 262K CHAR.	2,096K	7	8	32	.203	425,450	
7	Honeywell	1800	11/63	24.00	48	2.0	8 to 65K	3,120K	6	6	48	.438	1,365,840	
8	Honeywell	200/1200	1/66	21.67	32	1.5	16 to 131K CHAR.	786K	11	6	--	.346	272,160	
9	IBM	360/50	8/65	16.00	32	2.0	65 to 524K CHAR.	4,192K	3	8	--	.227	950,110	
10	Burroughs	3500	6/67	16.00	16	1.0	5 to 500K CHAR.	4,000K	4	8	16	.200	801,360	
11	Honeywell	200/200	7/64	16.00	32	2.0	2 to 65K CHAR.	390K	15	6	--	.499	184,600	
12	NCR	315	5/62	15.00	12	.8	10 to 80K	960K	10	4-Numeric 6-Alpha	12	.594	570,000	
13	UNIVAC	9300(2)	10/67	13.34	8	.6	8 to 32K CHAR.	256K	16	8	32	.405	103,780	Thin Film Memory
14	Burroughs	5500	11/64	12.00	48	4.0	32K	1,536K	9	6	48	.520	788,260	
15	RCA	70/45	7/66	11.15	16	1.44	16 to 262K CHAR.	2,096K	7	8	32	.248	523,800	

TABLE XIX

PERFORMANCE PARAMETERS FOR SELECTED BUSINESS-ORIENTED COMPUTERS (Continued)

Entry No.	Company	Model No.	Date 1st Installation	Bits per Micro-Second	Bits Per Cycle	Central Processor		Maximum Bit Capacity	Rank By Maximum Bit Capacity	Word Length		Cost Per Bit for Maximum Memory Capacity	Cost for Maximum Memory Capacity	Remarks
						Cycle Time (μs)	Number of Words			Bits/ Character	Bits/ Word			
16	RCA	70/35	10/66	11.15	16	1.44	16 to 65K CHAR.	520K	13	8	32	.391	203,700	16 Words of Nonaddressable Scratch Memory
17	Honeywell	200/120	2/66	10.67	16	1.5	2 to 32K CHAR.	192K	17	6	--	.570	109,500	
18	RCA	70/25	1/66	10.61	16	1.5	16 to 65K CHAR.	520K	13	8	32	.368	191,600	
19	General Electric	435	9/65	8.89	24	2.7	16 to 32K	768K	12	6	24	.461	354,240	
20	General Electric	437	---	8.89	24	2.7	16 to 131K	3,144K	5	6	24	.289	907,650	
21	Honeywell	800	12/60	8.00	48	6.0	4 to 32K	1,536K	9	6	48	.432	663,600	
22	Burroughs	2500	5/67	8.00	16	2.0	5 to 60K CHAR.	480K	14	8	16	.365	174,960	
23	RCA	3301	7/64	8.00	12	1.5	40 to 160K CHAR.	960K	10	6	--	.469	450,000	
24	UNIVAC	9200(2)	5/67	6.67	8	1.2	8 to 16K CHAR.	128K	18	8	32	.343	43,940	Thin Film Memory
25	IBM	360/40	5/65	6.40	16	2.5	16 to 262K CHAR.	2,096K	7	8	--	.232	485,290	
26	General Electric	425	6/64	6.15	24	3.9	8 to 32K	768K	12	6	24	.335	257,000	
27	General Electric	427	---	6.15	24	3.9	16 to 131K	3,144K	5	6	24	.250	786,550	
28	IBM	360/30	6/65	5.33	8	1.5	8 to 65K CHAR.	520K	13	8	--	.343	178,390	
29	General Electric	415	5/64	4.15	24	5.8	8 to 32K	768K	12	6	24	.267	205,420	
30	RCA	70/15	10/65	4.00	8	2.0	4 to 8K CHAR.	64K	20	8	32	.742	47,500	
31	General Electric	235	4/64	3.33	20	6.0	4 to 16K	96K	19	6	18	.303	87,120	
32	NCR	315/100	5/63	2.00	12	6.0	10 to 40K	480K	14	4-Numeric 6-Alpha	12	.716	344,000	

(1) Prices not available due to manufacturer pulling processors back for modifications.

(2) Total price includes printer.

(3) Total cost includes 500K words of disc storage. Cost per bit does not take into account this disc storage capacity.

TABLE XX

PERFORMANCE CHARACTERISTICS FOR SELECTED SCIENTIFIC-ORIENTED COMPUTERS

Entry No.	Company	Model No.	Date of 1st Installation	Bits Per Micro Second	Central Processor				Rank by Maximum Bit Capacity	Word Length		Cost Per Bit For Maximum Memory Capacity	Cost For Maximum Memory Capacity	Remarks
					Bits Per Cycle	Cycle Time (μs)	Number of Words	Maximum Bit Capacity		Bits/ Character	Bits/ Word			
1	CDC	6800(1)	----	240.00	60	.25	32-131K	7,860K	4	6	60			
2	IBM	360/65	3/66	83.33	64	.75	131 to 1024K CHAR	8,192K	3	8	--	.265	$2,174,740	
3	IBM	360/67	7/66	83.33	64	.75	131 to 1024K CHAR	8,192K	3	8	--	.285	2,335,350	
4	IBM	360/75	4/66	83.33	64	.75	262 to 1024K CHAR	8,192K	3	8	--	.305	2,497,750	
5	Burroughs	7506(2)	----	80.00	48	.6	96 to 3145K CHAR	24,960K	1	8	48	.279	6,961,000	Thin Film Memory
6	Burroughs	6506(2)	----	80.00	48	.6	96 to 3145K CHAR	24,960K	1	8	48	.271	6,759,600	Thin Film Memory
7	General Electric	635(2)	6/65	72.00	72	1.0 (2 Words)	32 to 262K	9,432K	2	6	36	.237	2,236,800	
8	General Electric	645(1)	----	72.00	72	1.0 (2 Words)	32 to 262K	9,432K	2	6	36			
9	CDC	6400	5/66	60.00	60	1.0	131K	7,860K	4	6	60	.302	2,373,000	
10	CDC	6600	8/64	60.00	60	1.0	131K	7,860K	4	6	60	.465	3,675,000	
11	CDC	3800	5/64	54.80	48	.88	262K	6,288K	5	6	24	.568	3,570,000	
12	UNIVAC	1108 II(2)	12/65	48.00	36	.75	65 to 262K	9,432K	2	-	36	.264	2,476,800	
13	UNIVAC	494	2/66	40.00	30	.75	16 to 131K	3,930K	7	-	30	.326	1,279,400	
14	Burroughs	7504(2)	----	40.00	48	1.2	96 to 3145K CHAR	24,960K	1	8	48	.212	5,288,000	
15	Burroughs	6504(2)	1/68	40.00	48	1.2	96 to 3145K CHAR	24,960K	1	8	48	.204	5,095,200	
16	Scientific Data	Sigma 5	12/67	37.75	32	.85	4 to 131K	4,192K	6	8	32	.203	850,700	
17	Scientific Data	Sigma 7	12/66	37.75	32	.85	4 to 131K	4,192K	6	8	32	.231	969,700	
18	General Electric	625(2)	4/65	36.00	72	2.0 (2 Words)	32 to 262K	9,432K	2	6	36	.218	2,060,800	
19	Digital Equipment	PDP-10	9/67	36.00	36	1.0	8 to 262K	9,432K	2	Variable	36			
20	Raytheon	520	10/65	24.00	24	1.0	4 to 32K	768K	13	6	24			
21	EAI	8400	7/65	18.33	32	1.75	8 to 65K	2,340K	9	-	36			
22	CDC	3100	12/64	13.71	24	1.75	4 to 32K	768K	13	6	24	.356	273,500	
23	Scientific Data	SDS 940	4/66	13.71	24	1.75	65K	1,560K	11	6	24			
24	Scientific Control	6700	9/67	13.71	24	1.75	32 to 131K	3,144K	8	6	24			
25	UNIVAC	418	8/63	9.00	18	2.0	16 to 65K	1,170K	12	-	18	.232	271,800	
26	UNIVAC	491	9/65	6.25	30	4.8	16 to 65K	1,950K	10	-	30	.321	625,000	
27	IBM	1130(3)	9/65	4.45	16	3.6	4 to 32K	512K	14	8	16	.265	135,450	

(1) Prices not available due to manufacturer processors for modifications.
(2) Multi-processor computer. Cost of only 1st processor included.
(3) Total cost includes 1 million bytes of disc storage. Cost per bit does not take into account this disc storage capacity.

7.0 DEVELOPMENT PROGRAM

During the course of this study, the rudimentary requirements for the data archives - or better said, a data center - to service the data collected from the Apollo Applications Program have been defined and studied. It has been unquestionably demonstrated that there is a keen user interest in AAP data thereby necessitating the creation of an active data center rather than a mere long term bulk storage facility. Because of the limited amount of published information from NASA about the particulars of the AAP (list and definition of approved experiments, number of missions, expected data, etc.), it has been difficult to get a true assessment of users' requirements; however, the interest generated in the types of experiments anticipated for the AAP was very strong, especially among the community of extractive industry. Time limitations of this contract have restricted the contact with potential users largely to representatives of extractive industries. Thus, one of the key tasks remaining is to expand potential user exposure to the fields of biomedicine and life support systems, communications, the physical sciences and space engineering. More will be said about this in Section 7.2.

Apart from motivating user interest to develop user requirements, system specifications were generated for both a full-blown "archives" system as well as an initial system. The creation and growth of the initial system will be extremely critical to the eventual success of the full-blown system, thus much more time and effort must be spent in defining the goals and evolutionary procedures of the initial system.

Finally, much continued effort is required in the area of system design, to develop performance specifications and delineate all software requirements. This problem should also be addressed.

7.1 Initial System

Much has been said in Section 3.0 about the mechanics and structure of the initial system over the first few years of operation. The question which remains, however, is how to best implement the system to provide for the most effective system growth. At this stage in time, three possibilities appear to exist which are as follows:

- Six to eight months prior to the first AAP flight, initiate a pilot storage and retrieval system to handle data from the Earth Resources program being conducted by MSC.
- Six to eight months prior to the first AAP flight, initiate a pilot storage and retrieval system to handle data from experiments flown during the Gemini program which would be of particular interest to users.

- Initiate the initial system with the inception of the first AAP flight.

Other possibilities may also exist, and some time should be spent to explore them. In any case, it is believed that some sort of pilot operation, i.e., operation before the first AAP flight, would be advantageous in that it would serve to establish procedures, develop the confidence of users, and demonstrate the need for both the AAPA and also the experiments to be flown during the AAP.

7.1.1 Aircraft Sensor Data. - Since 1964, MSC has been flying Earth Resources Aircraft Missions with increasing sophistication. Many of the sensors and sensor data evolving from this program are similar to some of those which can be expected to be employed during the AAP. Use of this data in the pilot system, therefore, is felt to be very prudent as it is expected that the Earth resources data will be among the most demanded of the AAP data. A list of the sensors employed in the aircraft missions along with the amount and output of data of 12, 24 and 36 missions a year (using a Convair 240 and Lockheed Electra NP3A) is provided below.

Sensor	Recording Media	Missions Per Year		
		12	24	36
RC-8	9 1/2 in. film; 75 ft/roll	550 ft/mo	1100 ft/mo	1650 ft/mo
Multiband Camers	70mm film; 150 ft/roll	600 ft/mo	1200 ft/mo	1800 ft/mo
AAS-5 (UV) Imager	35mm film; 75 ft/roll	50 ft/mo	100 ft/mo	150 ft/mo
Microwave Imager	70mm film; 75 ft/roll	75 ft/mo	150 ft/mo	225 ft/mo
RS-7 IR Imager	70mm film; 75 ft/roll	75 ft/mo	150 ft/mo	225 ft/mo
Side Looking Radar	5 in. film; 150 ft/roll	150 ft/mo	300 ft/mo	400 ft/mo
Dual Channel Imager	70mm film; 75 ft/roll	150 ft/mo	300 ft/mo	450 ft/mo
Recon IV IR Imager	70mm film; 75 ft/roll	75 ft/mo	150 ft/mo	225 ft/mo
KA62 Camera Cluster	5 in. film; 150 ft/roll	1800 ft/mo	3600 ft/mo	5400 ft/mo
Boresight Camera	35mm film; 1000 ft/roll	200 ft/mo	4000 ft/mo	6000 ft/mo

Sensor	Recording Media	Missions Per Year 12	24	36
Passive MW Radiometer	*Magnetic Tape	--	--	--
Scatterometer	Magnetic tape	--	--	--
IR Spectrometer	Magnetic tape	--	--	--
IR Radiometer	Magnetic tape	--	--	--
Multi-Frequency MW Radiometer	Magnetic tape	--	--	--
Laser Altimeter	Magnetic tape	--	--	--
PRT-5 (Barnes)	Magnetic tape	--	--	--
Total Air Temp Probe	Magnetic tape	--	--	--
ASQ-90	Magnetic tape	--	--	--
	Total Film	5,525 ft/mo 66,500 ft/yr	11,050 ft/mo 132,600 ft/yr	16,575 ft/mo 198,900 ft/mo
	Total Tape	7/mo 84/yr	14/mo 168/yr	21/mo 252/yr

* The magnetic tape recorded data can all be contained on the same tape during each mission.

The film to be used will consist of B&W, IR, color, color IR, and color aeroneg.

The magnetic tapes used will be one inch tapes and 5000 feet long.

Storage of this data in a pilot AAPA would thus serve to accomplish two functions, namely, 1) it would provide a good test of the utility of an AAPA (data center), and 2) it would enable procedures to be established and developed before the actual AAP data is available.

7.1.2 Gemini Data. - During the Gemini program, several experiments were performed which are similar to or forerunners of experiments to be performed during the AAP. Examples of this are experiment S005, Synoptic Terrain Photography, and experiment S006, Synoptic Weather Photography. Some of these experiments would have a very strong user appeal and could thus also serve as excellent material to store in a pilot archive. Needless to say, data from Gemini which is related to AAP experiments should be stored in the AAPA and more important could be used as a data base of a pilot system. A further examination of this possibility is recommended.

7.1.3 AAP Flight 1A Data. - The first scheduled AAP mission, designated AAP Flight 1A, is to be launched in mid-1969. Twenty-six experiments have been approved for the flight and are listed below.

NO.	NAME	MONITORING FACILITY	
D017	Carbon Dioxide Reduction	Huntsville	
S005	Synoptic Terrain Photography	Houston	E
S006	Synoptic Weather Photography	Houston	E
S009	Nuclear Emulsion	Houston	
S015	Zero-G Single Human Cells	Houston	
*S016	Trapped Particles Asymmetry	Houston	
S017	X-Ray Astronomy	Houston	
S018	Micro metoid Collection	Houston	
S019	Ultraviolet Stellar Astronomy	Houston	
S020	UV and X-ray Solar Photography	Houston	
*027	Galactic X-ray Astronomy	Huntsville	
S039	Day-Nite Camera	Houston	E
S043	IR Temperature Sounding	Houston	E

NO.	NAME	MONITORING FACILITY	
S049	IR Interferometer Spectrometer	Houston	E
S050	IR Temperature Profile Radiometer	Houston	E
S063	UV Air Glow Horizon Photography	Houston	
S065	Multiband Terrain Photography	Houston	
S075	Electric Scan Microwave Radiometer	Houston	E
S100	Metric Camera (Incl. Stellar)	Houston	E
S101	Multiband Photography	Houston	E
S102	Dual Channel Scanner Images	Houston	E
S103	Shortwave Length Spectrometer	Houston	E
S104	Microwave Temperature Sounder	Houston	E
T002	Manned Navigation Sighting	Houston	
T003	In-Flight Nephelometer	Houston	
T004	Frog Otolith Function	Houston	

In addition to these experiments, the following are being held as potential candidates.

NO.	NAME	MONITORING FACILITY	
S073	Geigenshine Zodiacal Light	Houston	
S105	Altimeter Scatterometer	Houston	E
S106	Radar Imager	Houston	E

* These items are questionable at the present time.

Those experiments marked by an "E" in the right-hand column are related to Earth resources and are therefore expected to generate much user interest. As may be seen, there are fourteen such experiments and a total data base of the order of 10^8 bits, to say nothing of photographic and film data.

With so many experiments having a potentially large user interest, it would be possible to establish the pilot AAPA on data from AAP 1-A alone. This would require organizing the system toward the end of the second quarter of 1969 to enable the personnel to establish procedures and arrange for the acquisition of ground truth data and data from experimenters prior to the flight. The structure of the pilot system, whether initiated toward the end of the second quarter of 1969 or shortly before (somewhere in the first quarter of 1969 would be recommended if Gemini or Aircraft Missions data is employed and the AAP 1-A date is met), would be as discussed in Sections 3.7 and 3.11. It is to be noted that whichever pilot technique is adopted, approximately a year will be required prior to its inception to establish all requirements - operational and procedural - and to properly acquaint a "bank" of users with the system through seminars and continued visitations as performed under this study.

7.2 User Investigation

As it has been a key element in this study, the interface with users will also be an essential feature in the development of the AAPA. For an undertaking such as the AAPA, it is extremely important to have an active body of users who are both aware of the system and the data stored therein, as well as the mechanics for extracting or retrieving the data. Thus, it is recommended that the user visitations and presentations be continued on an even larger scale to a broader base of industrial, academic, and governmental representatives. The presentations given to date have stimulated interest in Earth resources, sensors and sensor development, information concerning areas of the Earth where sensors will view, and new or modified experiments which could be performed. Thus, by delivering presentations to a larger base of potential users (covering more S.I.C. categories) and by utilizing a larger base of presentation materials, it is expected that a more complete description of users' needs could be gained and a firmer foundation established for exploitation of the AAP experiments data.

In addition to contacting a broader base of disciplines in the user survey, an effort should also be made to deliver presentations to a segment of small business. Experience gained from this study further indicates that in order to accommodate the small business activities with the AAPA, it will be necessary to provide a central repository with some peripheral equipment for analyzing data. Characteristically, small businesses do not have a large investment potential or much of the capital equipment which will be required to interpret much of the experimental data derived from the AAP. Thus, to accommodate small business and to provide for presentations to a broader base of disciplines, work should be performed to analyze their specific needs and to gather additional presentation material oriented to their needs.

The two outstanding problems which have been cited above are 1) the need to educate the broadest possible base of potential users and 2) the need to provide each user with sufficient information about the background of experiments to permit a self-analysis of data requirements. To help alleviate these problems, while at the same time providing a positive step toward the creation and organization of the AAPA, it is further recommended that NASA undertake to review and categorize (possibly under S.I.C. sector number designations) each experiment. That is, at some time during the cycle of approving an experiment, each experiment should be reviewed to determine its applicability to an economic (as opposed to scientific) endeavor and then to pinpoint the specific area. This would implement NASA's internal bookkeeping, initialize an AAPA reference index, and provide for efficient distribution of information to potential users.

7.3 System Design

In the area of system design, the first level functions of an AAPA system have been defined and explored, an estimate of the data base has been compiled and a review has been made of storage and retrieval devices and techniques. There is, however, a large gap to bridge before the actual implementation of the system is undertaken.

Perhaps the first thing required in this further investigation is a rigid definition of the scope of the system; i.e., what direction will the system take, what data will it store, and what data classes (experimenter tapes, reduced data tapes, raw data tapes, etc.) will it store. With this completed, the functional details behind the first level flow diagram (Section 3.7) may be established.

In Section 3.11, personnel considerations and a preliminary budget were delineated for the initial system. Further investigations will also have to be performed to refine this preliminary analysis and to examine the labor market to meet the demands of the system.

To complement the initial state of the art review, the hardware and software details must be defined so that procurement specifications can be generated. Associative memory techniques should be examined as a possible means of reducing the access problem and areas of further research in equipment may be defined where a developmental breakthrough might alleviate a hardware or software deficiency. Also, during this phase of hardware and software development, the personnel who will staff the pilot system should be brought together, including the data analysis, surrogation, and key management personnel, to investigate and create an accession code for all the data.

Finally, the feasibility of interorganizational relationships with agencies other than the NSSDC should be examined and all system interfaces established. The most critical of these interfaces will be that between the AAPA, NASA, and the experimenters. Current feelings are that the AAP staff should have direct contact with experimenters after going through the NASA field office responsible for the experiment. This, however, bears further investigation.

APPENDIX A

DIGITAL DATA ESTIMATES

FOR AAP

TABLE A-1
ESTIMATED DATA FIGURES FOR FIRST 192 EXPERIMENTS
(EXTRAPOLATED FROM DATA OBTAINED FOR FIRST 55 EXPERIMENTS)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
I. SPACE SCIENCE												
A. Astronomy												
S001 Zodiacal Light Photography		X										
S013 UV Astronomical Camera		X										
S017 X-Ray Astronomy		X		AAP-1A	5.8×10^7				X			
S019 UV Stellar Astronomy		X		AAP-1A	3.8×10^6		134			Slides		10,000
S027 Galactic X-Ray Mapping		X		AAP-1	3.5×10^7				X			
S030 Dim Sky Photography/or Thicon		X										
S069 X-Ray Astronomy (B)		X		AAP-2	5.8×10^7				X			
Group Subtotal					15.5×10^7		134					10,000
Percent of Scheduled Group							50					
Extrapolated Group Total					31.0×10^7		268					20,000
B. Planetary Atmosphere												
S011 Airglow Horizon Photography	X											
S021 Airglow Horizon Photography	X	X										
S025 IR Temperature Soundings	X											
S063 UV Airglow Horizon Photography	X	X		AAP-1A	1.2×10^5							40
Group Subtotal					1.2×10^5							40
Percent of Scheduled Group							25					
Extrapolated Group Total					4.8×10^5							160
C. Interplanetary Dust												
S020 UV X-Ray Solar Photography		X		AAP-1A	1.7×10^5							
S035 Medium Energy Solar Wind		X										
S038 Low Energy Solar Wind		X										
S064 UV Photography of Dust Clouds		X										
S070 UV/X-Ray Solar Photography												
T015 Meteoroid Composition		X										
T016 Meteoroid Entry Observation		X										
T017 Meteoroid Impact + Erosion		X		AAP-2								
T021 Meteoroid Velocity		X		AAP-2	1.0×10^9				X			
Group Subtotal					1.0×10^9							
Percent of Scheduled Group							22.2					
Extrapolated Group Total					4.5×10^9							

TABLE A-1 (Continued)

EXPERIMENT CATEGORIES & LISTINGS	TENTATIVE ARCHIVE AAPA	TENTATIVE ARCHIVE NSSDC	TENTATIVE ARCHIVE OTHER	SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD: ANALOG CHANNEL HOURS MEDIUM BAND	ANALOG CHANNEL HOURS WIDE BAND	FILM 8MM	FILM 16MM	FILM 35MM	FILM 70/75MM	PHOTOS
D. Planetology												
S031 Lunar Passive Seismology		X										
S032 Lunar Gravity		X										
S033 Lunar Active Seismology		X										
S034 Lunar Tri-Axis Magnetometer		X										
S036 Suprathermal ION Detection		X										
S037 Lunar Heat Flow		X										
S058 Cold Cathode Ionization Gauge		X										
S059 Lunar Geology Investigation		X										
Group Subtotal												
Percent of Scheduled Group					0							
Extrapolated Group Total												
II. PHYSICAL SCIENCES												
A. Solar Physics												
M454 Shadow Shield												
S010 Agena Micrometeorite Collect		X										
S012 Micrometeorite Collection												
S018 Micrometerorite Collection		X		AAP-1A								
S029 Libration Regions Photographs		X										
S052 White Light Coronagraph		X		AAP-4	6.3×10^8					X		
S053 UV Coronal Spectographs		X		AAP-4	9.7×10^8						X	
S054 X-Ray Spectrographic Telescope		X		AAP-4	2.7×10^8					X		
S055 UV Spectrometers		X		AAP-4	3.5×10^9						(X)	
S056 Dual X-Ray Telescopes		X		AAP-4	1.2×10^9					X		
S062 Solar X-Ray Flux Variation		X										
S067 Gamma Ray and X-Ray Spectroscopy		X										
Group Subtotal					6.8×10^9							
Percent of Scheduled Group					41.7							
Extrapolated Group Total					16.3×10^9							
B. Particles and Fields												
S009 Nuclear Emulsion		X		AAP-1A	1.9×10^5							
S016 Trapped Particles Asymmetry		X		AAP-1A								
S022 Low Z Cosmic Rays		X										
S023 High Z Cosmic Rays		X										
S024 Nuclear Emulsion		X										

TABLE A-1 (Continued)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
B. Particles and Field (Continued)												
S051 Daytime Sodium Cloud		X										
D008 Radiation in Spacecraft	X											
					1.9×10^5							
							16.7					
					11.4×10^5							
C. Ionospheric Physics												
S026 ION Wake Measurement		X										
							0					
III. SPACE ENGINEERING												
A. Engineering Activities												
M401 Mapping and Survey System	X			AAP-1A	2.8×10^8					X	(M)	
M402 Orbital Workshop	X			AAP-2							(M)	
M403 Electrostatic Charge	X											
M404 Proton/Electron		X										
M405 Tri-Axis Magnetometer		X										
M406 Optical Communication		X										
M407 Lunar UV Spectral Reflect		X										
M408 Beta Spectrometer		X										
M409 Bremsstrahlung Spectrometer		X										
M410 Color Patch Photography	X											
M411 2-Color Earth S Limb Photos	X											
M412 Landmark Contrast Meas	X											
M413 Subcritical Cryo Storage	X											
M415 Thermal Control Coatings	X											
M416 Propell and Mass Determination	X											
M417 Liquid Interface Stability	X											
M418 Boiling Heat Transfer	X											
M419 Cryogenic Propellant Transfer	X											
M420 Propell and Storage System	X											
M422 DC Motor + Gear Lubrication	X											

TABLE A-1 (Continued)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
A. Engineering Activities (Continued)												
M423 Hydrostatic Gas Bearing	X											
M426 Condensing Heat Transfer	X											
M427 Strapdown Platform	X											
M432 Large Space Structures	X											
M433 Satellite Recovery	X											
M436 Artifical Gravity		X										
M466 Space Suit Evaluation	X			AAP-2	1.5×10^8				X			
M469 ST-124 Removal	X			AAP-2	4.4×10^6				X			
M475 Water + Waste Management	X											
M484 Orbital Workshop Artificial G	X											
M486 Astronaut Eva Equipment	X											
M487 Habitability/Crew Quarters	X			AAP-2					X			
M492 Tube Joining in Space	X			AAP-2				X				
M493 Electron Beam Welding	X			AAP-2							(X)	
M495 Lunar Surface Engr Properties	X											
M496 Liquid Drop Dynamics	X											
M497 Fluid Density Gradient	X											
M508 Astronaut EVA Hardware Evaluation	X			AAP-2							(X)	
M509 Astronaut Maneuvering Evaluation	X			AAP-2							(X)	
1002 Moderate Depth Drill	X											
1004 Lunar Surveying System	X											
1010 Variable Gravity Support	X											
T019 Pegasus Panel Retrieval	X											
T020 Jet-Shoes	X			AAP-2					X			
D012 Astronaut Maneuvering Unit	X											
D016 Power Tool Evaluation	X											
D017 Carbon Dioxide Reduction	X			AAP-1A	1.2×10^6							
D018 Integrated Maintenance	X			AAP-2							(X)	
D019 Suit Donning + Sleep Sta Eval	X			AAP-2							(X)	
D020 Alternate Restraints Eval	X			AAP-2							(X)	
					4.4×10^8							
							8					
					55.0×10^8							

TABLE A-1 (Continued)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
B. Space Vehicle Technology												
M479 Zero Gravity Flammability	X			AAP-2	3.8×10^3				X			
M483 Radar Attitude Sensing System	X											
M488 High Pressure Gas Expulsion	X			AAP-2	5.0×10^5				X			
M489 Heat Exchanger	X			AAP-2	8.0×10^6						(X)	
T005 Fusible Material Radiator	X											
T022 Heat Pipe	X			AAP-1	4.6×10^5							
T023 Surface Absorbed Materials	X			AAP-2								
D003 Mass Determination	X											
D021 Expandable Airlock Technology	X			AAP-2	3.0×10^6				X	X		
D022 Expandable Reentry Structures	X			AAP-2	2.5×10^7	3-10			X			
					3.7×10^7							
						60						
					6.2×10^7							
C. Electronics Controls												
T003 In-Flight Nephelometer	X			AAP-1A								
T011 Reentry Communications	X											
T013 Crew-Vehicle Disturbance	X											
D002 Nearby Object Photography	X											
						0						
D. Space Application												
D001 Basic Object Photography	X											
D006 Surface Photography	X											
D007 Space Object Radiometry	X											
D015 Night Image Intensification	X											
						0						

TABLE A-1 (Continued)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
E. Navigation and Traffic												
M439 Star Horizon Automatic Tracking		X		AAP-3	7.2×10^4							
M441 Noncooperative Tracking Radar	X											
T002 Manual Navigation Sightings	X			AAP-1A	1.1×10^5							
T014 Orbital Horizon Definition	X											
D004 Celestial Radiometry		X										
D005 Star Occultation Navigation	X											
D009 Simple Navigation	X											
D010 ION-Sensing Attitude Control	X											
					1.8×10^5							
							25					
					7.2×10^5							
F. Lunar and Planetary SC												
1001 Simple Navigation	X											
1009 Nuclear Emulsion	X											
							0					
IV. LIFE SCIENCES												
A. Biomedical and Behavioral												
M001 Cardiovascular Conditioning	X											
M003 In-Flight Exerciser	X											
M004 In-Flight Phonocardiogram	X											
M005 Bioassays Body Fluids	X											
M006 Bone Demineralization	X											
M007 Calcium Balance Study	X											
M008 In-Flight Sleep Analysis	X											
M009 Human Otolith Function	X											
M011 Cytogenetic Blood Studies	X											
M012 Exercise Ergometer	X											
M017 Thoracic Blood Flow	X											
M018 Vectorcardiogram	X			AAP-2,4	2.3×10^8							
M019 Metabolic Rate Measurement	X											
M020 Pulmonary Function	X											
M021 Semicircular Canal Function	X											

TABLE A-1 (Continued)

EXPERIMENT CATEGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
A. Biomedical and Behavioral (Continued)												
M022 Red Blood Cell Survival	X											
M023 Lower Body Negative Pressure	X											
M048 Anti G Elastic Garment	X											
M049 Analysis CO + CO_2	X											
M050 Metabolic Activity	X			AAP-2,4	2.0×10^8							
M051 Cardiovascular Function Assess	X			AAP-2,4	1.8×10^8							
M052 Bone and Muscle Changes	X			AAP-1,2,3,4								
M053 Human Vestibular Function	X			AAP-2,4	5.0×10^5						(X)	
M054 Neurological Study (EEG)		X			4.1×10^9							
M055 Time and Motion	X			AAP-2,4					X			
T006 Vision Test Equip Eval	X											
T007 Human Transfer Functions	X											
D013 Astronaut Visibility	X											
					4.7×10^9							
							17.9					
					26.2×10^9							
B. Environment Biology												
S002 Sea Urchin Egg Growth	X											
S004 Radiation + Zero G on Blood	X											
S061 Potato Respiration	X			AAP-3	6.7×10^5							
T009 Primates in Long Term Zero Gravity	X											
					6.7×10^5							
							25					
					26.8×10^5							
C. Physical Biology												
S008 Visual Acuity	X											
S015 Zero G Single	X			AAP-1A					X			

TABLE A-1 (Continued)

	EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
							ANALOG CHANNEL HOURS		FILM				
		AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
	D. Biotechnology												
	T004 Frog Otolith Function	X			AAP-1A	3.2×10^7	10 Hrs.						
	T008 Electrolysis Cell	X											
	T010 On-Board Centrifuge	X											
						3.2×10^7							
						10.7×10^7		33.3					
V.	COMMUNICATIONS												
	A. Communications												
	M430 LM Relay	X											
	M440 Wide Band Variable Power Trans	X											
	1003 Emplaced Scientific Station	X											
	S041 Millimeter Wave Propagation	X											
	T001 Reentry Communications	X											
	T012 Optical Technology	X											
	D014 UHF-VHF Polarization		X										
								0					
VI.	METEOROLOGY												
	A. Meteorology												
	S006 Synoptic Weather Photography	X			AAP-1A							(X)	
	S007 Cloud Top Spectrometer		X										
	S039 Day-Night Camera System		X										
	S040 Dielectric Tape Camera System	X											
	S043 IR Temperature Sounding		X										
	S044 02 + H20 Microwave Radiometer		X										
	S045 IR Filter Wedge Spectrometer		X										
	S046 Visible RAD Polarization Meas	X											
	S047 Stellar Refract Density Meas	X											
	S048 UHF Sferics Detection	X											
	S049 IR Interferometer Spectrometer		X										

TABLE A-1 (Continued)

EXPERIMENT CATOGORIES & LISTINGS	TENTATIVE ARCHIVE			SCHEDULED LAUNCH ASSIGNMENT	DIGITAL (BITS)	DESIRED DATA YIELD						
						ANALOG CHANNEL HOURS		FILM				
	AAPA	NSSDC	OTHER			MEDIUM BAND	WIDE BAND	8MM	16MM	35MM	70/75MM	PHOTOS
A. Meteorology (Continued)												
S050 15 Micron Grating Spectrometer		X										
S057 Multi-Channel Radiometer		X										
S060 Selective Chopper Radiometer		X										
							0					
VII. EARTH RESOURCES												
A. Natural Resources												
S005 Synoptic Terraine Photography	X			AAP-1A								40
S042 Multi-Spectral Photography	X	X										
S063 Multi-Bank Terrain Photography	X	X		AAP-1A, 2, 3							(X)	
							0					

TABLE A-2
DIGITAL DATA ESTIMATES FOR AAP

EXPERIMENT	EXTRAPOLATED DATA TOTALS FOR 192 EXPTS. *	DISTRIBUTION BETWEEN NSSDC	DISTRIBUTION BETWEEN AAPA
I SPACE SCIENCE			
A. Astronomy	3.1×10^8 Bits	3.1×10^8	
B. Planetary Atmosphere	4.8×10^5	2.4×10^5	4.8×10^5
C. Interplanetary Dust	4.5×10^9	4.5×10^9	
D. Planetology	Not Available	All	
Factored Subtotal	6.4×10^9	6.4×10^9	4.8×10^5
II PHYSICAL SCIENCES			
A. Solar Physics	1.6×10^{10}	1.6×10^{10}	
B. Particles and Fields	1.1×10^6	1.1×10^6	
C. Ionospheric Physics	Not Available	All	
Factored Subtotal	2.4×10^{10}	2.4×10^{10}	
III SPACE ENGINEERING			
A. Engineering Activity	5.5×10^9	8×10^8	4.7×10^9
B. Space Vehicle Technology	6.2×10^7		6.2×10^7
C. Electrical Controls	Not Available		All
D. Space Application	Not Available		All
E. Navigation and Traffic	7.2×10^5	1.8×10^5	5.4×10^5
F. Lunar and Planetology SC	Not Available		All
Factored Subtotal	1.1×10^{10}	8×10^8	1.1×10^{10}
IV LIFE SCIENCES			
A. Biomedical and Behavior	2.6×10^{10}		2.6×10^{10}
B. Environment Biology	2.7×10^6		2.7×10^6
C. Physical Biology	Not Available		All
D. Biotechnology	1.1×10^8		1.1×10^8
Factored Subtotal	3.5×10^{10}		3.5×10^{10}
V COMMUNICATIONS	Not Available	As Desired	All
VI METEOROLOGY	Not Available	As Desired	All
VII EARTH RESOURCES	Not Available	As Desired	All
Total for 192 Expts.		3.1×10^{10}	4.6×10^{10}
2.4 "Best-Fit" Factor		7.5×10^{10}	10.5×10^{10}
400/192 Factor		1.6×10^{11}	2.2×10^{11}

*See Rpt. Figure for Method

Combined Total 3.8×10^{11} Bits

APPENDIX B

SUMMARY OF TRIPS

Date	Trip	Pertinent Information
5/19	C. W. Williams (CWW) to NASA/ERC	Review of documentation; Review Test project plans
5/23	CWW to NASA/ERC	MIT presentation; Boston Edison (Mr. Sweeney): Conference calls to University of New Mexico and Indiana Memorial Union; Meeting with J. Dennison-NASA Technology Utilization
5/24	Dr. Stanley Ross and Mr. Maurice Moroney to UAC	Organization meeting ; D.T.G. Photostore; Demonstration of hybrid system
5/25	CWW to Washington	NSSDC (James Fava and John Cambell)
5/26		GSFC (Instrumentation Analysis - R. Holmes and J. J. Kearney)
5/31		NASA/Headquarters (Dr. Fordyce, Mr. Krueger)
6/1	CWW to Washington	National Archives (Dr. Aldrich; Bellcomm, Inc., (Phillip Schaenman)
6/2		NSSDC (John Cambell)
6/5	CWW to Ithaca	Cornell University (Art McNair, TaLiang)
6/6	CWW to New York	Contact (Mr. T. Harriman, VP); NYC Department of Air Pollution (Dr. Ferrand)

Date	Trip	Pertinent Information
6/7	G. E. Schmidt, Jr. and D. L. Sharp Washington	GSFC (Telemetry and Data Acquisition - Bill Anonsen)
6/14	Mr. Moroney to UAC	Delivery of documentation; plan IBM trip; discussion of I.T.S. problem
6/22	CWW to Kollsman Instrument Corporation	Explanation of UAC mission in AAP Data Archives and make arrangements for formal presentation
6/27	CWW to Fort Belvoir, Virginia	Obtain information concerning AAP data related to Mapping and Geodecy
6/29	G. E. Schmidt, Jr. and CWW to NASA/ERC	Discussion of previsit package and presentation
7/7	G. E. Schmidt, Jr. and CWW to NASA/ERC	Review of first monthly progress report and material for previsit package and presentation
7/20-21	CWW to NASA Headquarters, Washington, D.C., and Fort Belvoir, Virginia	NASA/Hdq, - Investigation of present NASA interface with the National Archives System and the Federal Records Center
		Acquisition of information about some specific experiments from Dr. Jocelyn Gill
		NASA/Hdq, Mr. Ted George - Discussion of acquisition of the Bellcomm estimate of the AAP data base and the acquisition of copies of NASA forms 1138 and 1038 for several experiments

(Continued on next page)

Date	Trip	Pertinent Information
7/20-21		NASA/Hdq, Mr. Sam Fodyce - Review of the AAP presentation for potential users; discussion of NASA's experiments review board Fort Belvoir, Dr. Colvocorresses - Discussion of specific experiments and digital analysis of pictorial data
		NASA/Hdq, Mr. Spriggs - Acquisition of names of any industrial representatives on NASA's experiments review board
7/26	CWW to NASA/ERC	Review of AAP Presentation to be given to potential users
7/27	CWW to NASA/MSC, Houston, Texas	Investigation of NASA Photographic Library
7/28	CWW to University of New Mexico	Investigation of operations of participant in NASA's Technology Utilization Group
8/3	CWW and D. L. Sharp to Bloomfield, Connecticut	Presentation to Connecticut General Life Insurance Company
8/15	G. E. Schmidt, Jr. and D. L. Sharp to Greenbelt, Maryland	Investigation of NSSDC
8/23	CWW to NASA/ERC	Review of present plan, contract administration, and budget

Date	Trip	Pertinent Information
8/24	CWW and G. E. Schmidt, Jr. to New York City	Presentation to Kollsman Instrument Corporation
		Presentation to New York City Department of Air Pollution Control
8/28	CWW to Tacoma, Washington	Presentation to Weyerhaeuser Company
8/29	CWW to Salt Lake City, Utah	Presentation to Kennecott Copper Corporation
8/30	CWW to San Jose, California	Presentation to IBM Corporation and examine IBM trillion bit memory
8/31	CWW to Redwood City, California	Presentation to Ampex Corporation
9/1	CWW to Moffett Field, California	Presentation to NASA/Ames Research Center
9/19	CWW to New York City	Presentation to Dr. E. I. Piore, IBM Corporation
9/20	CWW to New York City	Presentation to Shell Oil Company
10/3	CWW to New York City	Presentation to New York City Department of Air Pollution Control and discussion of potential pollution space experiment

Date	Trip	Pertinent Information
10/4	CWW to New York City	Presentation to Shell Development Company
10/10	CWW to Bethlehem, Pennsylvania	Presentation to Bethlehem Steel Corporation
10/16	CWW to Pittsburg, Pennsylvania	Presentation to United States Steel Corporation
10/18	CWW to NASA/ERC	Discussion of AAP flight IA
10/24	CWW to NASA/ERC	Discussion of AAP flight IA
11/7	CWW to NASA/ERC	Discussions with S. Fordyce, M. Maroney, S. Ross, and C. Huebner about flight IA
12/12	CWW to NASA/MSC	Discussions about flight IA and the Aircraft Earth Resources Missions
12/13	CWW to NASA/MSFC	Discussions to flight IA and review of Ampex Videofile and Recordac Miracode systems used at Huntsville
12/20	CWW Williams to NASA/ERC	Discussions about flight IA and review of MSC requirements for flight IA
1/8	CWW Williams to NASA/ERC	Discussions with G. Larson, M. Moroney, and S. Ross concerning the Earth Resources Program at ERC

APPENDIX C

BIBLIOGRAPHY

Fanwick, C., Trends in Computer Hardware. System Development Corporation, March 1966.

"Integrated Computer Memories." Scientific American.

"A Summary of Progress in the Computer Field." Computer Design, October 1966.

"Where is Computer Speed Going?." Computer Design, October 1966.

"Programming Languages." Computer Design, April 1966.

Schneidewind, N. F., "The Practice of Computer Selection." Datamation, February 1967.

Herman, D. J., "SCERT: A Computer Evaluation Tool." Datamation, February 1967.

"Precision Magnetic Tape." Datamation, October 1966.

"Changes in Computer Performance." Datamation, September 1966.

Hobbs, L. C., "The Impact of Hardware in the 1970's." Datamation, March 1966.

Siera, M. M., "Mass Memories." WESCON/65, August 1965.

Weniger, K., "Memory Systems Comparison." The Electronic Engineer, May 1967.

Ringer, M., "Mass Storage Devices: What's in Store?." Computer Design, December 1966.

Craver, J. S., "A Review of Electromechanical Mass Storage." Datamation, July 1966.

"Superconductive Memories After 11 Years." Engineering Design News, 1967.

Thermoplastic Recording: A Year of Progress. Quarterly Electronics Digest, General Electric Company.

Kuehler and Kerby, A Photo-Digital Mass Storage System. 1966 Fall Joint Computer Conference.

"Computer Industry Annual 1967-1968." Compiled by the Editors of Computer Design.

Gianala, V. F., Plated Wire Magnetic Film Memories. Bell Telephone Labs, December 19

Casewell, H. L., Cryoelectronics. AIME Transaction, Vol. 236, March 1966.

Hanlon, A. G., Content-Addressable and Associative Memory Systems - A Survey. IEEE Transactions on Electronic Computers, Vol. EC-15, August 1966.

Beiser, L., Laser Beam and Electron Beam Extremely Wideband Information Storage and Retrieval. Photographic Science and Engineering, Vol. 10, July - August 1966.

Boysel, L., "Memory on a Chip - A Step Towards Large-Scale Integration." Electronics, Vol. 40, February 1967.

Lynch, T. J., Space Data Handling at Goddard Space Flight Center. NASA-TM-X-55180, February 1965.

Chong, C., Medium Speed Mass Random - Access Memory Final Report. UNIVAC Engineering Center/Griffiss AFB, March 1965.

Barboga, G., Summary of Efforts Expended on the Concept of Dynamic Update of a Microfilm File. MITRE Corporation, AFSC/ESD, August 1965.

Dlugatch, I., Large Capacity Laser Memory for Spaceborne Computers. 1st Annual Princeton Conference on Information Science and Systems, March 1967.

Lowe, T. C., Design Principles of an On-Line Information Retrieval System. December 1966.

Braddock, J. V., Dove Data Storage and Retrieval System. March 1967,

Price, T. J., Non-Destructive Read Only Memories. Memorandum Report, Frankford Arsenal, May 1966.

Frate, G. J., Compendium of Storage and Retrieval Devices and Techniques. RADC, September 1964.

Scientific American Information Issue, September 1966.

Sammet, J. E., "Programming Languages: Current and Future Trends." Computers and Automation, March 1967.

Radocy, F., "Tape Storage Problems." Journal of the Audio Engineering Society, January 1957.

Santoro, J., Compendium of 70 mm x 100 mm Unit Record Production Viewing and Storage-Retrieval Equipment. RADC-TR-66-359 Final Report, September 1966.

Gagliandi, U. O., Data File Size and Its Relation to the Bayesian Effectiveness of an Information Retrieval System. ESD-TR-65-275, April 1965.

Bervl, L., Information Storage and Retrieval: A State-of-the-Art Report. Auerback Corporation, September 1964.

Ringer, M., "Mass Storage Devices: Whats In Store?" Computer Design, December 1966.

Whitaker, L. A., "Exploring the Mystery of Magnetic Tape." DPMA Quarterly, October 1964.

Mullen, J. T., "Is Magnetic Tape Long Lived?" Audio, June 1965.

Westmijze, W. K., Studies on Magnetic Tape Recording. Philips Research Report, June 1953.

Rogers, K. T. et al, High Information Density Storage Devices. Stanford Research Institute Technical Report Ecom-01261-3, June 1966.

Brandhorst, W. T., and Eckert, P.F., Guide to the Processing, Storage, and Retrieval of Bibliographic Information at the NASA Scientific and Technical Information Facility. June 1966.

Weygand, A. G., Communications and Tracking for First Four Missions of Apollo Applications Program. Bellcomm Inc. Technical Memorandum TM-67-2034-I, February 16, 1967.

Gadberry, H. M., Utilization of NASA-Generated Space Technology by Midwestern Industry. Midwest Research Institute, Quarterly Progress Report No. 1, February 23, 1962, CR-5064

Kimball, Dr. Charles N., Opportunities for the Practical Use of Space Technology by Businessmen. CR-60784.

Burnett, James E., The Utilization of Technological Resources. NASA-TM X-52248, January 1966.

Gadberry, H. M., Utilization of NASA-Generated Space Technology by Midwestern Industry. Quarterly Progress Report No. 3, CR-50648, August 1962.

Seamans, Dr. Robert C., Jr., Address before the New York Board of Trade, N66-39948, October 27, 1966.

Hirsch, Werner Z., Transformation of New Knowledge for Economic Growth. UCLA MR-8 June 1964.

Apollo Applications Program. Technical Summary, NASA-OMSF, May 1, 1967.

Telemetry Data Processing. NASA-GSFC, X-564-65-384, September 1965.

Guidelines for Principal Investigators Submitting Data to the National Space Science Data Center. NASA-NSSDC, July 1966.

Satellites and Sounding Rockets. Goddard Projects Summary, 1964.

OGO Program Bibliography. NASA-NSSDC, October 1966.

Radocy, Frank, "Trap Storage Problems." Journal of the Audio Engineering Society, January 1957.

Development and Application of Computer Software Techniques to Human Factors Task Data Handling Problems. AMRL-TR-66-200, Joint NASA/USAF Study, December 1966.

"FY 1967 to Bring Hardware Starts." Missiles and Rockets, November 29, 1965, pp.48-54.

Habit, Edmund J., Keipert, Frank A., and Lee, Richard C., Telemetry Processing for NASA Scientific Satellites. Instrument Society of America Preprint July 22, 1965.

"Data Processing Plan for Orbiting Solar Observatory (OSO-E)." NASA-GFSC, X-565-66345, August 1966.

Holmes, Richard G., Poland, William B., Demmerle, Allan M., The GSFC Scientific Data Storage Problem. NASA-GSFC, X-563-67-82, January 1967.

Opportunities for Participating in Space Flight Investigations. NASA-DSSA, NHB 8030.1, July 1965

Allen, T. J., The Differential Performance of Information Channels in the Transfer of Technology. Sloan School of Management, MIT, June 1966.

Goodman, A. F. et al, DOD User - Needs Study, Phase II: Flow of Scientific and Technical Information within the Defense Industry. Autonetics Division, North American Aviation, 3 vols (AD 647111; AD 647112; AD 649284), November 1966.

Allen, T. J., Managing the Flow of Scientific and Technological Information. Sloan School of Management, MIT, (Doctoral disseration), September 1966.

Bervl, Lawrence, Information Storage and Retrieval; a State-of-the-Art Report. Auerback Corporation, (AD 630089), September 1964.

Kent, Allen, Textbook on Mechanized Information Retrieval. 2nd ed. John Wiley & Sons, New York, 1966.

Meadow, Charles T., The Analysis of Information Systems. John Wiley & Sons, New York, 1967.

Wasserman, Paul, and Bundy, Mary Lee, Manpower for the Library and Information Professions in the 1970's; an Inquire into Fundamental Problems. School of Library and Information Services, University of Maryland, September 1966.

Garvey, W. D., and Griffith, Belver C., "Scientific Communication as a Social System." Science, v157 (1967), pp. 1011-1016.

APPENDIX D

NEW TECHNOLOGY

No reportable inventions or other innovations, improvements, or discoveries which may have more than academic significance were made or conceived during the performance of this contract.